# Decomposition Optimization Algorithms for Distributed Radar Systems

Ying Ma, Sheng Chen, *Fellow, IEEE*, Chengwen Xing, *Member, IEEE*, Xiangyuan Bu, and Lajos Hanzo, *Fellow, IEEE*



***Abstract*—Distributed radar systems are capable of enhancing the detection performance by using multiple widely spaced distributed antennas. With prior statistic information of targets, resource allocation is of critical importance for further improving the system's achievable performance. In this paper, the total transmitted power is minimized at a given mean-square target-estimation error. We derive two iterative decomposition algorithms for solving this nonconvex constrained optimization problem, namely, the optimality condition decomposition (OCD)-based and the alternating direction method of multipliers (ADMM)-based algorithms. Both the convergence performance and the computational complexity of our algorithms are analyzed theoretically, which are then confirmed by our simulation results. The OCD method imposes a much lower computational burden per iteration, while the ADMM method exhibits a higher per-iteration complexity, but as a benefit of its significantly faster convergence speed, it requires less iterations. Therefore, the ADMM imposes a lower total complexity than the OCD. The results also show that both of our schemes outperform the state-of-the-art benchmark scheme for the multiple-target case, in terms of the total power allocated, at the cost of some degradation in localization accuracy. For the single-target case, all the three algorithms achieve similar performance. Our ADMM algorithm has similar total computational complexity per iteration and convergence speed to those of the benchmark.**

***Index Terms*—Alternating direction method of multipliers, localization, multiple-input multiple-output radar, optimality condition decomposition, resource allocation.**

## I. Introduction

MULTIPLE-input multi-output (MIMO) radar systems relying on widely-separated antennas have attracted considerable attention from both industry and academia. The family of distributed MIMO radar systems is capable of significantly improving the estimation/detection performance [1]–[6] by exploiting the increased degrees of freedom resulting from the improved spatial diversity. In particular, distributed radar systems are capable of improving accuracy of target location and velocity estimation by exploiting the different Doppler estimates from multiple spatial directions [7]–[10].

Naturally, the localization performance of MIMO radar systems relying on widely-spaced distributed antennas, quantified in terms of the mean square estimation error (MSE), is determined by diverse factors, including effective signal bandwidth, the signal-to-noise ratio (SNR), the product of the numbers of transmit and receive antennas, etc [11]. Since the SNR is influenced by the path loss, the target radar cross section (RCS) and the transmitted power, the attainable localization performance can be improved by increasing either the number of participating radars or the transmitted power. However, simply increasing the amount of resources without considering the cooperation among the individual terminals is usually far from optimal.

In most traditional designs, the system's power budget is usually allocated to the transmit radars and it is fixed [6], [10], which is easy to implement and results in the simplest network structure. However, when prior estimation of the target RCS is available, according to estimation theory, uniform power allocation is far from the best strategy. In battlefields, a radar system is usually supported by power-supply trucks, but under hostile environments, their number is strictly limited. Thus, how to allocate limited resources to multiple radar stations is of great importance for maximizing the achievable performance. In other words, power allocation substantially affects the detection performance of multi-radar systems.

Recently, various studies used the Cramer-Rao lower bound (CRLB) for evaluating the performance of MIMO radar systems [11]–[16]. A power allocation scheme [12] based on CRLB was designed for multiple radar systems with a single target. The resultant nonconvex optimization problem was solved by relaxation and a domain-decomposition method. Specifically, in [12] the total transmitted power was minimized at a given estimation MSE threshold. However the algorithm of [12] was not designed for multiple-target scenarios, which are often encountered in practice. In [13] a power allocation algorithm was proposed for the multiple-target case, which is equally applicable to the single-target senario.

Against this background, in this paper, we propose two iterative decomposition methods, which are referred to as the optimality condition decomposition (OCD) [17] and the alternating direction method of multipliers (ADMM) [18] algorithms, in order to minimize the total transmitted power while satisfying a predefined estimation MSE threshold. These two algorithms can be applied to both multiple-target and single-target scenarios. The ADMM method has been widely adopted for solving convex problems. In this paper, we extend the ADMM algorithm to nonconvex problems and show that it is capable of converging.

Manuscript received March 8, 2016; revised June 17, 2016 and July 18, 2016; accepted August 9, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Fauzia Ahmad. This work was supported in part by the National Natural Science Foundation of China under Grants 61421001 and 61671058.

Y. Ma, C. Xing, and X. Bu are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: mayingbit2011@gmail.com; xingchengwen@gmail.com; bxy@bit.edu.cn).

S. Chen is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K., and also with King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: sqc@ecs.soton.ac.uk).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2016.2602801

It is worth pointing out that Simonetto and Leus [19] applied the ADMM method to solve a localization problem in a sensor network by converting the nonconvex problem to a convex one using rank-relaxation. However, the algorithm of [19] cannot be applied to our problem, because the task of [19] is that of locating sensors, which is not directly related to the signal waveform and power. Furthermore, the maximum likelihood (ML) criterion can be used for solving this sensor localization problem. However, our task is to assign the power of every MIMO radar transmitter, and at the time of writing it is an open challenge to design the ML estimator for this task [11]. The main contributions of our work are as follows.

- We propose two iterative decomposition algorithms, namely, the OCD-based and ADMM-based methods, for both multiple-target and single-target scenarios. The convergence of these two algorithms is analyzed theoretically and verified by simulations. Both these two methods are capable of converging to locally optimal solutions. The complexity analysis of the two algorithms is provided and it is shown that the OCD method imposes a much lower computational burden per iteration, while the ADMM method enjoys a significantly faster convergence speed and therefore it actually imposes a lower total complexity.
- In the multiple-target case, we demonstrate that both of our two algorithms outperform the state-of-the-art benchmark scheme of [13], in terms of the total power allocated at the expense of some degradation in localization accuracy. We show furthermore that our ADMM-based algorithm and the algorithm of [13] have similar convergence speed and total computational complexity.
- In the single-target case, we show that all the three methods attain a similar performance, since the underlying optimization problems are identical. We also prove that the closed-form solution of [12] is invalid for the systems with more than three transmit radars and we propose a beneficial suboptimal closed-form solution.

The paper is organized as follows. In Section II, the MIMO radar system model is introduced and the corresponding optimization problem is formulated. Our power allocation strategies are proposed in Section III for both the multiple-target and single-target cases, while our convergence and complexity analysis is provided in Section IV. Section V presents our simulation results for characterizing the attainable performance of the proposed algorithms which are then compared to the scheme of [13]. Finally, our conclusions are offered in Section VI.

Throughout our discussions, the following notational conventions are used. Boldface lower- and upper-case letters denote vectors and matrices, respectively. The transpose, conjugate and inverse operators are denoted by $(\cdot)^{\mathrm{T}}$, $(\cdot)^{*}$ and $(\cdot)^{-1}$, respectively, while $\mathrm{Tr}\,(\cdot)$ stands for the matrix trace operation and $\mathrm{diag}\,(x_1, x_2, \cdots, x_n)$ or $\mathrm{diag}(\mathbf{x})$ is the diagonal matrix with the specified diagonal elements. Additionally, $\mathrm{diag}\,(\mathbf{X}_1, \cdots, \mathbf{X}_K)$ and $\mathrm{diag}\,(\mathbf{x}_1, \cdots, \mathbf{x}_K)$ denotes the block diagonal matrices with the specified sub-matrices and vectors, respectively, at the corresponding block diagonal positions. The operator $\mathrm{v}_{\mathrm{diag}}(\mathbf{X})$ forms a vector using the diagonal elements of square matrix $\mathbf{X}$, while $\mathrm{E}\{\cdot\}$ denotes the expectation operator and $\otimes$ is the Kronecker product operator. The sub-matrix consisting of the elements of the $i_1$ to $i_2$ rows and $j_1$ to $j_2$ columns of $\mathbf{A}$ is denoted by $[\mathbf{A}]_{[i_1:i_2;j_1:j_2]}$, and the $i$th row and $j$th column element of $\mathbf{A}$ is given by $[\mathbf{A}]_{i,j}$. Similarly, $[\mathbf{a}]_{[i_1:i_2]}$ is the vector consisting of $i_1$th to $i_2$th elements of $\mathbf{a}$. The magnitude operator is given by $|\cdot|$, and $\|\cdot\|$ denotes the vector two-norm or matrix Frobenius norm. $\mathbf{I}_K$ is the identity matrix of size $K \times K$ and $\mathbf{0}$ is the zero matrix/vector of an appropriate size, while $\mathbf{1}$ denotes the vector of an appropriate size, whose elements are all equal to one. Finally, $\Re[\,]$ denotes the real part of a complex value and $\mathrm{j} = \sqrt{-1}$ represents the imaginary axis.

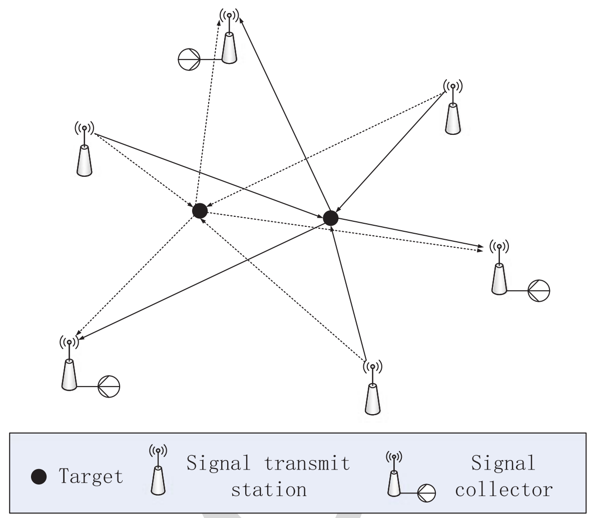



Fig. 1. Illustration of distributed radar network.

## II. System Model

The MIMO radar system consists of $M$ transmit radars and $N$ receive radars which cooperate to locate $K$ targets, as illustrated in Fig. 1. The $M$ transmit radars are positioned at the coordinates $(x_m^{tx}, y_m^{tx})$ for $1 \le m \le M$, and the $N$ receive radars are positioned at $(x_n^{rx}, y_n^{rx})$ for $1 \le n \le N$, while the position of target $k$ is $(x_k, y_k)$. A set of mutually orthogonal waveforms are transmitted from the transmit radars, and the corresponding baseband signals are denoted by $\left\{s_m(t)\right\}_{m=1}^{M}$ with normalized power, i.e., $\int_{\tau_m} |s_m(t)|^2\, dt = 1$, where $\tau_m$ is the duration of the $m$th transmitted signal. Furthermore, the orthogonality of the transmitted waveforms can always be guaranteed even for different time delays, i.e., $\int_{\tau_m} s_m(t) s_{m'}^{*}(t-\tau)\, dt = 0$ for $m' \neq m$. The narrowband signals of the transmitted waveforms have the effective bandwidth $\beta_m$ specified by

$$\beta_m^2 = \frac{\int_W f^2 \left|S_m(f)\right|^2\, df}{\int_W \left|S_m(f)\right|^2\, df}\,(\mathrm{Hz})^2, \tag{1}$$

where $W$ is the frequency range of the signals, and $S_m(f)$ is the Fourier transform of $s_m(t)$ transmitted from the $m$th transmit radar. The transmitted powers of the different antennas, denoted by $\mathbf{p} = [p_1\, p_2 \cdots p_M]^{\mathrm{T}}$, are constrained by their corresponding minimum and maximum values specified by

$$\mathbf{p}_{\min} = \left[p_{1_{\min}}\, p_{2_{\min}} \cdots p_{M_{\min}}\right]^{\mathrm{T}}, \tag{2}$$

$$\mathbf{p}_{\max} = \left[p_{1_{\max}}\, p_{2_{\max}} \cdots p_{M_{\max}}\right]^{\mathrm{T}}. \tag{3}$$

The upper bound $p_{m_{\max}}$ is determined by the design and the lower bound $p_{m_{\min}}$ is used to guarantee that the transmit radar $m$ operates at an appropriate SNR. Let the propagation path spanning from the transmitter $m$ to the target $k$ and from the target $k$ to the receiver $n$ be defined as the channel $(m,k,n)$. Then the propagation time $\tau_{m,n}^{(k}$ of the channel $(m,k,n)$ can be calculated by $\tau_{m,n}^{(k} = (R_{m,k}^{tx} + R_{n,k}^{rx})/c$, where $c$ is the speed of light, $R_{m,k}^{tx} = \sqrt{(x_m^{tx} - x_k)^2 + (y_m^{tx} - y_k)^2}$ is the distance from transmitter $m$ to target $k$, and $R_{n,k}^{rx} = \sqrt{(x_n^{rx} - x_k)^2 + (y_n^{rx} - y_k)^2}$ is the distance from target $k$ to receiver $n$. The time delay $\tau_{m,n}^{(k}$ is used to estimate the position of targets. For far field signals, by retaining only the linear terms of its Taylor expansion, $\tau_{m,n}^{(k}$ can be approximated as a linear function of $x_k$ and $y_k$

$$\tau_{m,n}^{(k} \simeq -\frac{x_k}{c}\left(\cos\theta_m^{(k} + \cos\varphi_n^{(k}\right) - \frac{y_k}{c}\left(\sin\theta_m^{(k} + \sin\varphi_n^{(k}\right), \tag{4}$$

where $\theta_m^{(k}$ is the bearing angle of the transmitting radar $m$ to the target $k$ and $\varphi_n^{(k}$ is the bearing angle of the receiving radar $n$ to the target $k$, both measured with respect to the $x$ axis.

Let the complex-valued reflectivity coefficient $h_{m,n}^{(k}$ represent the attenuation and phase rotation of channel $(m,k,n)$. The baseband signal at receive radar $n$ can be expressed as

$$r_n(t) = \sum_{k=1}^{K}\sum_{m=1}^{M} \sqrt{p_m}\, h_{m,n}^{(k} s_m\left(t - \tau_{m,n}^{(k}\right) + \omega_n(t), \tag{5}$$

where $\omega_n(t)$ is a circularly symmetric complex Gaussian white noise, which is bandlimited to the system bandwidth $W$ and hence has a zero mean and $E\{|\omega_n(t)|^2\} = \sigma^2$. In our work, the path-loss $\kappa_{m,n}^{(k}$ is chosen as

$$\kappa_{m,n}^{(k} \propto \frac{1}{\left(R_{m,k}^{tx}\right)^2\left(R_{n,k}^{rx}\right)^2}. \tag{6}$$

Thus, given the complex target RCS $\zeta_{m,n}^{(k}$, the channel coefficient $h_{m,n}^{(k}$ is given by

$$h_{m,n}^{(k} = \zeta_{m,n}^{(k}\sqrt{\kappa_{m,n}^{(k}} = h_{m,n}^{(k,\mathrm{Re}} + \mathrm{j}h_{m,n}^{(k,\mathrm{Im}}, \tag{7}$$

where $h_{m,n}^{(k,\mathrm{Re}}$ and $h_{m,n}^{(k,\mathrm{Im}}$ are the real and imaginary parts of $h_{m,n}^{(k}$. Let us collect all the channel coefficients associated with the target $k$ in the $(2MN \times 1)$-element real-valued vector as

$$\mathbf{h}_k = \left[h_{1,1}^{(k,\mathrm{Re}} \cdots h_{1,N}^{(k,\mathrm{Re}} \cdots h_{M,N}^{(k,\mathrm{Re}}\, h_{1,1}^{(k,\mathrm{Im}} \cdots h_{1,N}^{(k,\mathrm{Im}} \cdots h_{M,N}^{(k,\mathrm{Im}}\right]^{\mathrm{T}}. \tag{8}$$

Similarly, we introduce the $(NM \times 1)$-element real vectors

$$\left|\mathbf{h}^{(k}\right|^2 = \left[\left|h_{1,1}^{(k}\right|^2 \cdots \left|h_{1,N}^{(k}\right|^2 \cdots \left|h_{M,1}^{(k}\right|^2 \cdots \left|h_{M,N}^{(k}\right|^2\right]^{\mathrm{T}}, \tag{9}$$

$$\left|\mathbf{h}^{(k}\right| = \left[\left|h_{1,1}^{(k}\right| \cdots \left|h_{1,N}^{(k}\right| \cdots \left|h_{M,1}^{(k}\right| \cdots \left|h_{M,N}^{(k}\right|\right]^{\mathrm{T}}. \tag{10}$$

Upon defining $\mathbf{h} = \left[\mathbf{h}_1^{\mathrm{T}}\, \mathbf{h}_2^{\mathrm{T}} \cdots \mathbf{h}_K^{\mathrm{T}}\right]^{\mathrm{T}}$ and the location vector of the $K$ targets as $\boldsymbol{l}_{x,y} = \left[x_1\, y_1 \cdots x_K\, y_K\right]^{\mathrm{T}}$, all the system's parameters can be stacked into a single real-valued vector

$$\mathbf{u} = \left[\boldsymbol{l}_{x,y}^{\mathrm{T}}\, \mathbf{h}^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{11}$$

Since the received signal (5) is also a function of the time delays $\tau_{m,n}^{(k}$, we also define the following system parameter vector

$$\boldsymbol{\psi} = \left[\tau_{1,1}^{(1} \cdots \tau_{1,N}^{(1} \cdots \tau_{M,N}^{(K}\, \mathbf{h}^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{12}$$

There exists a clear one-to-one relationship between $\mathbf{u}$ and $\boldsymbol{\psi}$.

Let $f(\mathbf{r}|\mathbf{u})$ be the conditional probability density function (PDF) of the observation vector $\mathbf{r} = [r_1(t), r_2(t), \cdots, r_N(t)]$ conditioned on $\mathbf{u}$. Similarly, we have the conditional PDF of $\mathbf{r}$ conditioned on $\boldsymbol{\psi}$. Then the unbiased estimate $\widehat{\mathbf{u}}$ of $\mathbf{u}$ satisfies the following inequality [20]

$$\mathrm{E}\left\{\left(\widehat{\mathbf{u}} - \mathbf{u}\right)\left(\widehat{\mathbf{u}} - \mathbf{u}\right)^{\mathrm{T}}\right\} \geq \mathbf{J}^{-1}(\mathbf{u}), \tag{13}$$

where the Fisher information matrix (FIM) $\mathbf{J}(\mathbf{u})$ is defined by

$$\mathbf{J}(\mathbf{u}) = \mathrm{E}\left\{\frac{\partial}{\partial\mathbf{u}}\log f(\mathbf{r}|\mathbf{u})\left(\frac{\partial}{\partial\mathbf{u}}\log f(\mathbf{r}|\mathbf{u})\right)^{\mathrm{T}}\right\}. \tag{14}$$

Similarly, we have the FIM of $\boldsymbol{\psi}$, denoted by $\mathbf{J}(\boldsymbol{\psi})$. The FIM $\mathbf{J}(\mathbf{u})$ can be derived from $\mathbf{J}(\boldsymbol{\psi})$ according to

$$\mathbf{J}(\mathbf{u}) = \begin{bmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{2KMN}\end{bmatrix}\mathbf{J}(\boldsymbol{\psi})\begin{bmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{2KMN}\end{bmatrix}^{\mathrm{T}}, \tag{15}$$

where the $(2K \times KMN)$-element block diagonal matrix $\mathbf{D}$ takes the following form

$$\mathbf{D} = \mathrm{diag}\left(\mathbf{D}^{(1}, \mathbf{D}^{(2}, \cdots, \mathbf{D}^{(K}\right), \tag{16}$$

with the $(2 \times MN)$-element sub-matrix $\mathbf{D}^{(k}$ given by

$$\mathbf{D}^{(k} = \begin{bmatrix}\dfrac{\partial\tau_{1,1}^{(k}}{\partial x_k} & \cdots & \dfrac{\partial\tau_{M,N}^{(k}}{\partial x_k}\\ \dfrac{\partial\tau_{1,1}^{(k}}{\partial y_k} & \cdots & \dfrac{\partial\tau_{M,N}^{(k}}{\partial y_k}\end{bmatrix} = -\frac{1}{c}\begin{bmatrix}\cos\left(\theta_1^{(k)}\right)+\cos\left(\varphi_1^{(k)}\right) & \cdots & \cos\left(\theta_M^{(k)}\right)+\cos\left(\varphi_N^{(k)}\right)\\ \sin\left(\theta_1^{(k)}\right)+\sin\left(\varphi_1^{(k)}\right) & \cdots & \sin\left(\theta_M^{(k)}\right)+\sin\left(\varphi_N^{(k)}\right)\end{bmatrix}. \tag{17}$$

The matrix $\mathbf{C}_{x,y}$ associated with the CRLB for the unbiased estimator of $\boldsymbol{l}_{x,y}$ is the $(2K \times 2K)$-element upper left block sub-matrix of $\mathbf{J}^{-1}(\mathbf{u})$, which can be derived as [11], [21]

$$\mathbf{C}_{x,y} = \left[\mathbf{J}^{-1}(\mathbf{u})\right]_{[1:2K;1:2K]} = \left(\mathbf{D}\mathbf{P}\boldsymbol{\Psi}\mathbf{D}^{\mathrm{T}}\right)^{-1}, \tag{18}$$

where $\mathbf{P} = \mathbf{I}_K \otimes \mathrm{diag}(\mathbf{p}) \otimes \mathbf{I}_N$, and $\boldsymbol{\Psi} = \mathrm{diag}\big(\boldsymbol{\Psi}^{(1}, \cdots, \boldsymbol{\Psi}^{(K)}\big)$ is the $(KMN \times KMN)$-element block diagonal matrix with the $k$th sub-matrix defined as

$$\boldsymbol{\Psi}^{(k} = 8\pi^2 \left(\mathrm{diag}\left(\beta_1^2, \cdots, \beta_M^2\right) \otimes \mathbf{I}_N\right) \mathrm{diag}\left(\left|\mathbf{h}^{(k}\right|^2\right). \quad (19)$$

Let us denote the variances of the estimates of $x_k$ and $y_k$ by $\sigma_{x_k}^2$ and $\sigma_{y_k}^2$, respectively. Then we have

$$\sum_{k=1}^{K} \left(\sigma_{x_k}^2 + \sigma_{y_k}^2\right) \geq \mathrm{Tr}\left(\mathbf{C}_{x,y}\right), \quad (20)$$

where $\mathrm{Tr}\left(\mathbf{C}_{x,y}\right)$ is a lower bound on the sum of the MSEs of the localization estimator $\widehat{\boldsymbol{l}}_{x,y}$. By defining $\mathbf{X} = \mathrm{diag}(\mathbf{p}) \otimes \mathbf{I}_N$ and noting $\mathbf{D}$ of (16), we obtain the expression of the lower bound for the $k$th target location estimate as [12], [22]

$$\begin{aligned} &\sum_{i=1}^{2} \left[\mathbf{C}_{x,y}\right]_{i+2(k-1),i+2(k-1)} \\ &= \sum_{i=1}^{2} \left[\left(\mathbf{D}\mathbf{P}\boldsymbol{\Psi}\mathbf{D}^{\mathrm{T}}\right)^{-1}\right]_{i+2(k-1),i+2(k-1)} \\ &= \mathrm{Tr}\left(\begin{bmatrix} \left(\mathbf{a}_{1,1}^{(k}\right)^{\mathrm{T}} \mathbf{p} & \left(\mathbf{a}_{1,2}^{(k}\right)^{\mathrm{T}} \mathbf{p} \\ \left(\mathbf{a}_{2,1}^{(k}\right)^{\mathrm{T}} \mathbf{p} & \left(\mathbf{a}_{2,2}^{(k}\right)^{\mathrm{T}} \mathbf{p} \end{bmatrix}^{-1}\right) = \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}}{\mathbf{p}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}}, \quad (21) \end{aligned}$$

where the second equation is obtained by first dividing the $(MN \times 2)$ matrix $\left(\mathbf{D}^{(k)}\right)^{\mathrm{T}}$ into the two column vectors, $\left(\mathbf{D}^{(k)}\right)^{\mathrm{T}} = \left[\mathbf{d}_1^{(k} \, \mathbf{d}_2^{(k}\right]$, and generating the $(N \times 1)$ vectors

$$\mathbf{d}_{i,m}^{(k} = \left[\mathbf{d}_i^{(k}\right]_{[(m-1)N+1:mN]}, \; i = 1, 2, \; 1 \leq m \leq M. \quad (22)$$

Then $\mathbf{a}_{i,j}^{(k}$ for $1 \leq i, j \leq 2$ are given by

$$\begin{aligned} \mathbf{a}_{i,j}^{(k} = \mathrm{v}_{\mathrm{diag}} &\left(\mathrm{diag}\left(\left(\mathbf{d}_{i,1}^{(k}\right)^{\mathrm{T}}, \cdots, \left(\mathbf{d}_{i,M}^{(k}\right)^{\mathrm{T}}\right) \boldsymbol{\Psi}^{(k}\right. \\ &\left. \times \mathrm{diag}\left(\mathbf{d}_{j,1}^{(k}, \cdots, \mathbf{d}_{j,M}^{(k}\right)\right), \quad (23) \end{aligned}$$

while $\mathbf{b}_k = \mathbf{a}_{1,1}^{(k} + \mathbf{a}_{2,2}^{(k}$ and $\mathbf{A}_k = \mathbf{a}_{1,1}^{(k} \left(\mathbf{a}_{2,2}^{(k}\right)^{\mathrm{T}} - \mathbf{a}_{1,2}^{(k} \left(\mathbf{a}_{2,1}^{(k}\right)^{\mathrm{T}}$.

Our task is to design a beneficial power allocation strategy capable of achieving a localization accuracy threshold $\eta$. We can use the weighting $v_k$ to indicate the localization accuracy requirement for the $k$th target. The larger $v_k$ is, the higher accuracy is required for the $k$th target. For a predetermined lower bound of total MSE of all the targets, the transmit power of the different transmit radars can then be appropriately allocated for minimizing the total transmit power. This can be formulated as the following optimization problem $\mathbb{P}1$

$$\mathbb{P}1: \begin{array}{ll} \min\limits_{\mathbf{p}} & \mathbf{1}^{\mathrm{T}} \mathbf{p}, \\ \text{s.t.} & \sum\limits_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}}{\mathbf{p}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}} \leq \eta, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \quad (24)$$

Because generally speaking $\mathbf{A}_k$ is not a positive definite matrix, the optimization $\mathbb{P}1$ is a nonconvex problem.

In [13], a similar optimization problem is formulated as

$$\begin{array}{ll} \min\limits_{\mathbf{p}} & \mathbf{1}^{\mathrm{T}} \mathbf{p}, \\ \text{s.t.} & \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}}{\mathbf{p}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}} \leq \bar{\eta}, \; 1 \leq k \leq K, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M, \end{array} \quad (25)$$

given an equivalent localization accuracy threshold $\bar{\eta}$. In [13], a Taylor series based technique is applied to approximate the inequality constraints in (25) in order to relax the nonconvex optimization problem for the sake of obtaining a solution. Intuitively, the cost function associated with an optimal solution of our problem $\mathbb{P}1$ of (24) is generally smaller than that associated with an optimal solution of (25), i.e., we can achieve a lower power consumption. This is achieved at the potential cost of a slightly reduced localization accuracy.

## III. Power Resource Allocation

### A. Multi-Target Case

In order to solve the nonconvex problem $\mathbb{P}1$ of (24), we have to change it into a simpler form. Specifically, we have to change the inequality constraint into an equality one, i.e.,

$$\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}}{\mathbf{p}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}} \leq \eta \Rightarrow \sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}}{\mathbf{p}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}} = \eta. \quad (26)$$

*Lemma 1:* An increase of the transmit power $\mathbf{p}$ results in a reduction of the MSE, namely,

$$\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} (\mathbf{p} + \Delta\mathbf{p})}{(\mathbf{p} + \Delta\mathbf{p})^{\mathrm{T}} \mathbf{A}_k (\mathbf{p} + \Delta\mathbf{p})} \leq \sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}}{\mathbf{p}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}}. \quad (27)$$

The proof of Lemma 1 is similar to that of single-target case given in [12]. Thus, to achieve a reduced power consumption, we can always set the MSE to its maximum tolerance. The change of constraint as given in (26) leads to the problem $\mathbb{P}2$,

$$\mathbb{P}2: \begin{array}{ll} \min\limits_{\mathbf{p}} & \mathbf{1}^{\mathrm{T}} \mathbf{p}, \\ \text{s.t.} & \sum\limits_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}}{\mathbf{p}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}} = \eta, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \quad (28)$$

*Theorem 1:* The solutions of $\mathbb{P}1$ and $\mathbb{P}2$ are identical.

The proof of Theorem 1 is straightforward. By introducing the auxiliary variables

$$w_k = \frac{1}{\eta \mathbf{p}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}}, \; 1 \leq k \leq K, \quad (29)$$

and their corresponding lower and upper bounds

$$w_{k_{\min}} = \frac{1}{\eta \mathbf{p}_{\max}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}_{\max}}, \; w_{k_{\max}} = \frac{1}{\eta \mathbf{p}_{\min}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}_{\min}}, \; 1 \leq k \leq K, \quad (30)$$

$\mathbb{P}2$ is reformulated as the following optimization problem $\mathbb{P}3$:

$$\mathbb{P}3: \quad \begin{aligned} &\min_{\mathbf{p},\mathbf{w}} \ \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ &\text{s.t.} \ \sum_{k=1}^{K} v_k w_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} = 1, \\ &\qquad w_k \eta \mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} = 1,\ 1 \le k \le K, \\ &\qquad p_{m_{\min}} \le p_m \le p_{m_{\max}},\ 1 \le m \le M, \\ &\qquad w_{k_{\min}} \le w_k \le w_{k_{\max}},\ 1 \le k \le K. \end{aligned} \tag{31}$$

The following corollary is obvious.

*Corollary 1:* If $\mathbf{p}^\star$ associated with $w_k^\star = \frac{1}{\eta\left(\mathbf{p}^\star\right)^{\mathrm{T}}\mathbf{A}_k\mathbf{p}^\star}$ for $1 \le k \le K$ is an optimal solution of the problem $\mathbb{P}3$ (31), $\mathbf{p}^\star$ is an optimal solution for the problem $\mathbb{P}1$ of (24). Conversely, if $\mathbf{p}^\star$ is an optimal solution of the problem $\mathbb{P}1$, together with $w_k^\star = \frac{1}{\eta\left(\mathbf{p}^\star\right)^{\mathrm{T}}\mathbf{A}_k\mathbf{p}^\star}$ for $1 \le k \le K$ it is an optimal solution of the problem $\mathbb{P}3$.

*1) OCD-based method:* The Lagrangian associated with the optimization problem $\mathbb{P}3$ is

$$\begin{aligned} L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) = \mathbf{1}^{\mathrm{T}}\mathbf{p} + \lambda\left(\sum_{k=1}^{K} v_k w_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1\right) \\ + \sum_{k=1}^{K}\mu_k\left(w_k\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} - 1\right), \end{aligned} \tag{32}$$

with $\mathbf{w} = \left[w_1\, w_2 \cdots w_K\right]^{\mathrm{T}}$ and $\boldsymbol{\mu} = \left[\mu_1\, \mu_2 \cdots \mu_K\right]^{\mathrm{T}}$, where $\lambda$ and $\mu_k$ for $1 \le k \le K$ are Lagrangian multipliers. We optimize the Lagrangian (32) with respect to $\mathbf{p}$, $\lambda$, $w_k$ and $\mu_k$. Using the steepest descent method, the search directions are related to the Karush-Kuhn-Tucker (KKT) conditions by

$$\begin{aligned} \Delta\mathbf{p} = \nabla_{\mathbf{p}} L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) = \mathbf{1} + \lambda\left(\sum_{k=1}^{K} w_k v_k \mathbf{b}_k\right) \\ + \sum_{k=1}^{K}\mu_k w_k \eta\left(\mathbf{A}_k + \mathbf{A}_k^{\mathrm{T}}\right)\mathbf{p}, \end{aligned} \tag{33}$$

$$\Delta\lambda = \nabla_{-\lambda} L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) = -\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} + 1, \tag{34}$$

$$\begin{aligned} \Delta w_k &= \nabla_{w_k} L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) \\ &= \lambda v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} + \mu_k\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p},\ 1 \le k \le K, \end{aligned} \tag{35}$$

$$\begin{aligned} \Delta\mu_k &= \nabla_{-\mu_k} L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) \\ &= -\left(\eta w_k\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} + 1\right),\ 1 \le k \le K, \end{aligned} \tag{36}$$

where we have $\Delta\mathbf{p} = \left[\Delta p_1\, \Delta p_2 \cdots \Delta p_M\right]^{\mathrm{T}}$. The primal and dual variables are updated iteratively

$$p_m^{(n+1)} = \left[p_m^{(n)} - \kappa_1\Delta p_m^{(n)}\right]_{p_{m_{\min}}}^{p_{m_{\max}}},\ 1 \le m \le M, \tag{37}$$

$$\lambda^{(n+1)} = \lambda^{(n)} - \kappa_2\Delta\lambda^{(n)}, \tag{38}$$

$$w_k^{(n+1)} = w_k^{(n)} - \kappa_3\Delta w_k^{(n)},\ 1 \le k \le K, \tag{39}$$

$$\mu_k^{(n+1)} = \mu_k^{(n)} - \kappa_4\Delta\mu_k^{(n)},\ 1 \le k \le K, \tag{40}$$

where the superscript $^{(n)}$ denotes the iteration index and

$$[a]_b^c = \min\left\{\max\left\{a,b\right\},c\right\}, \tag{41}$$

while $\kappa_i$ for $1 \le i \le 4$ represent the step sizes for the primal variables $\mathbf{p}$, the dual variable $\lambda$, the primal variables $\mathbf{w}$ and the dual variables $\boldsymbol{\mu}$, respectively. According to [23], an exponentially decreasing step size is highly desired. Furthermore, since $\mathbf{p}$, $\lambda$, $\mathbf{w}$ and $\boldsymbol{\mu}$ have very different properties and their impacts on the Lagrangian are 'unequal', using different step sizes for them makes sense. By combining these two considerations, we set the four step sizes for $\mathbf{p}$, $\lambda$, $\mathbf{w}$ and $\boldsymbol{\mu}$ according to

$$\kappa_i = c_i e^{-\alpha n} \text{ with } 0 \le \alpha \ll 1, \text{ for } 1 \le i \le 4, \tag{42}$$

where $c_i > 0$ for $1 \le i \le 4$ are different constants.

The choice of the initial values for the primal variables $p_m$, $1 \le m \le M$, influences the convergence performance. Ideally, they should be chosen to be close to their own specific optimal values so as to enhance the convergence speed. For practical reason, the initialization should be easy and simple to realize too. Hence we opt for the initial powers of

$$\mathbf{p}^{(0)} = \mathbf{p}_{equ} = \frac{1}{\eta}\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{A}_k\mathbf{1}}\mathbf{1}, \tag{43}$$

which is obtained by setting all the elements of $\mathbf{p}$ to be equal. Then, $w_k$ is initialized according to

$$w_k^{(0)} = \frac{1}{\eta\mathbf{p}_{equ}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_{equ}},\ 1 \le k \le K. \tag{44}$$

The iterative procedure of (37) to (40) is repeated until $\left\|\mathbf{p}^{(n+1)} - \mathbf{p}^{(n)}\right\|$ becomes smaller than a preset small positive number or the maximum number of iterations is reached.

*Remark 1*: It is difficult to find a closed-form solution from the set of KKT conditions, because $\mathbf{A}_k$ for $1 \le k \le K$ are generally non-invertible. Hence our algorithm finds a locally optimal point in an iterative manner. It is also worth noting that the standard OCD [17] is typically based on a Newton-type algorithm, but our proposed OCD method is a steepest descent algorithm. The reason is that the Hessian matrix for the Lagrangian $L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu})$ of (32) is not invertible.

*2) ADMM-based method:* ADMM was originally proposed for solving convex problems in a parallel manner [18]. Let us now discuss how to apply the ADMM method for solving the nonconvex problem $\mathbb{P}3$. By introducing an auxiliary vector $\mathbf{z} = \mathbf{p}$, (29) can be rewritten as

$$\mathbf{p} = \mathbf{z} \ \text{ and } \ \eta w_k\mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} = 1,\ 1 \le k \le K. \tag{45}$$

Therefore, $\mathbb{P}3$ can be reformulated into the problem $\mathbb{P}4$:

$$\mathbb{P}4: \quad \begin{aligned} &\min_{\mathbf{p},\mathbf{w},\mathbf{z}} \ \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ &\text{s.t.} \ \sum_{k=1}^{K} v_k w_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} = 1, \\ &\qquad \mathbf{p} = \mathbf{z}, \\ &\qquad w_k \eta \mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} = 1,\ 1 \le k \le K, \\ &\qquad p_{m_{\min}} \le p_m \le p_{m_{\max}},\ 1 \le m \le M, \\ &\qquad w_{k_{\min}} \le w_k \le w_{k_{\max}},\ 1 \le k \le K. \end{aligned} \tag{46}$$

This problem is convex with respect to $\mathbf{p}$, $\mathbf{z}$ and $w_k$, respectively. An augmented Lagrangian is constructed as follows

$$L(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_0, d_1, \mathbf{d}_2) = \mathbf{1}^{\mathrm{T}}\mathbf{p} + \frac{\rho_0}{2}\|\mathbf{p}-\mathbf{z}\|^2 + \mathbf{d}_0^{\mathrm{T}}(\mathbf{p}-\mathbf{z}) + \sum_{k=1}^{K}\frac{\rho_{2,k}}{2}\left|w_k\mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta - 1\right|^2 + \sum_{k=1}^{K} d_{2,k}\left(w_k\mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta - 1\right) + \frac{\rho_1}{2}\left|\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1\right|^2 + d_1\left(\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1\right) \tag{47}$$

where $\mathbf{d}_0 = \left[d_{0,1}\cdots d_{0,M}\right]^{\mathrm{T}}$, $d_1$ and $\mathbf{d}_2 = \left[d_{2,1}\cdots d_{2,K}\right]^{\mathrm{T}}$ are the dual variables corresponding to the constraints $\mathbf{p} = \mathbf{z}$, $\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} = 1$ and $w_k\mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta = 1$ for $1 \le k \le K$, respectively, while $\rho_0$, $\rho_1$ and $\boldsymbol{\rho}_2 = \left[\rho_{2,1}\cdots\rho_{2,K}\right]^{\mathrm{T}}$ are the penalty parameters. Note that the augmented Lagrangian (47) is quadratic. For convenience, we scale the dual variables as $\mathbf{e} = \frac{1}{\rho_0}\mathbf{d}_0$, $\mu = \frac{1}{\rho_1}d_1$ and $\boldsymbol{\gamma} = \left[\gamma_1\cdots\gamma_K\right]^{\mathrm{T}}$ with $\gamma_k = \frac{1}{\rho_{2,k}}d_{2,k}$ for $1 \le k \le K$. Then, from (47) we obtain the following augmented Lagrangian

$$L(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{e}, \mu, \boldsymbol{\gamma}) = \mathbf{1}^{\mathrm{T}}\mathbf{p} + \frac{\rho_0}{2}\|\mathbf{p}-\mathbf{z}+\mathbf{e}\|^2 - \frac{\rho_0}{2}\|\mathbf{e}\|^2 + \sum_{k=1}^{K}\frac{\rho_{2,k}}{2}\left|w_k\mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta - 1 + \gamma_k\right|^2 - \sum_{k=1}^{K}\frac{\rho_{2,k}}{2}\left|\gamma_k\right|^2 + \frac{\rho_1}{2}\left|\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1 + \mu\right|^2 - \frac{\rho_1}{2}|\mu|^2. \tag{48}$$

We can find the saddle point of the augmented Lagrangian (48) by minimizing the Lagrangian over the primal variables $\mathbf{p}$, $\mathbf{w}$ and $\mathbf{z}$, as well as maximizing it over the dual variables $\mathbf{e}$, $\mu$ and $\boldsymbol{\gamma}$, in an alternative way. In particular, we update the primal variables $\mathbf{p}$, $\mathbf{w}$ and $\mathbf{z}$ separately one by one. Furthermore, after the update of the dual variables $\mathbf{e}$, $\mu$ and $\boldsymbol{\gamma}$, we adjust the penalty parameters $\rho_0$, $\rho_1$ and $\boldsymbol{\rho}_2$. We now summarize our ADMM-based procedure.

*Initialization:* Let us also opt for the equal power initialization $\mathbf{p}^{(0)} = \mathbf{p}_{equ}$ of (43). The other primal variables are initialized as $w_k^{(0)} = \frac{1}{\eta\mathbf{p}_{equ}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_{equ}}$ for $1 \le k \le K$ of (44), and

$$\mathbf{z}^{(0)} = \mathbf{p}_{equ}. \tag{49}$$

The initial penalty parameters, $\rho_0^{(0)}$, $\rho_1^{(0)}$ and $\rho_{2,k}^{(0)}$ for $1 \le k \le K$, are typically set to a large positive value, say, 500. Next, the dual variables are initialized as follows. Choose $\mu^{(0)} = 1$ and $\gamma_k^{(0)} = 1$ for $1 \le k \le K$, while every element of $\mathbf{e}^{(0)}$ is set to 1 too. Then we set the iteration index $n = 0$.

*Iterative Procedure:* At the $(n+1)$th iteration, perform:

- *Step 1: Update the primal variables* $\mathbf{p}$. Upon isolating all the terms involving $\mathbf{p}$ in the Lagrangian (48), we have

$$\min_{\mathbf{p}} \mathbf{1}^{\mathrm{T}}\mathbf{p} + \frac{\rho_0^{(n)}}{2}\left\|\mathbf{p}-\mathbf{z}^{(n)}+\mathbf{e}^{(n)}\right\|^2 + \frac{\rho_1^{(n)}}{2}\left|\sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1 + \mu^{(n)}\right|^2 + \sum_{k=1}^{K}\frac{\rho_{2,k}^{(n)}}{2}\left|w_k^{(n)}\left(\mathbf{z}^{(n)}\right)^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta - 1 + \gamma_k^{(n)}\right|^2,$$
$$\text{s.t.}\quad p_{m_{\min}} \le p_m \le p_{m_{\max}},\ 1 \le m \le M, \tag{50}$$

which is a constrained convex optimization. Setting the derivative of the objective function to zero yields the $(n+1)$th estimate of $\mathbf{p}$ as follows. First compute

$$\bar{\mathbf{p}}^{(n+1)} = \left[\bar{p}_1^{(n+1)}\cdots\bar{p}_M^{(n+1)}\right]^{\mathrm{T}} = \left(\mathbf{P}_1^{(n+1)}\right)^{-1}\mathbf{p}_2^{(n+1)}, \tag{51}$$

$$\mathbf{P}_1^{(n+1)} = \rho_0^{(n)}\mathbf{I}_M + \rho_1^{(n)}\left(\sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k\right) \times \left(\sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k^{\mathrm{T}}\right) + \sum_{k=1}^{K}\rho_{2,k}^{(n)} \times \left(w_k^{(n)}(\mathbf{A}_k)^{\mathrm{T}}\mathbf{z}^{(n)}\eta\right)\left(w_k^{(n)}\left(\mathbf{z}^{(n)}\right)^{\mathrm{T}}\mathbf{A}_k\eta\right)^{\mathrm{T}}, \tag{52}$$

$$\mathbf{p}_2^{(n+1)} = -\mathbf{1} + \rho_0^{(n)}\left(\mathbf{z}^{(n)} + \mathbf{e}^{(n)}\right) + \rho_1^{(n)}\left(\sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k\right)\left(1-\mu^{(n)}\right) + \rho_{2,k}^{(n)}\left(w_k^{(n)}(\mathbf{A}_k)^{\mathrm{T}}\mathbf{z}^{(n)}\eta\right)\left(1-\gamma_k^{(n)}\right). \tag{53}$$

The final estimate is then given by

$$p_m^{(n+1)} = \left[\bar{p}_m^{(n+1)}\right]_{p_{m_{\min}}}^{p_{m_{\max}}},\ 1 \le m \le M. \tag{54}$$

- *Step 2: Update the primal variables* $\mathbf{w}$. The optimization involving $\mathbf{w}$ is also a constrained convex problem

$$\min_{\mathbf{w}} \frac{\rho_1^{(n)}}{2}\left|\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p}^{(n+1)} - 1 + \mu^{(n)}\right|^2 + \sum_{k=1}^{K}\frac{\rho_{2,k}^{(n)}}{2}\left|w_k\left(\mathbf{z}^{(n)}\right)^{\mathrm{T}}\mathbf{A}_k\mathbf{p}^{(n+1)}\eta - 1 + \gamma_k^{(n)}\right|^2,$$
$$\text{s.t.}\quad w_{k_{\min}} \le w_k \le w_{k_{\max}},\ 1 \le k \le K. \tag{55}$$

The solution is given by

$$w_k^{(n+1)} = \left[\frac{w_{k,1}^{(n+1)}}{w_{k,2}^{(n+1)}}\right]_{w_{k_{\min}}}^{w_{k_{\max}}},\ 1 \le k \le K, \tag{56}$$

where

$$w_{k,1}^{(n+1)} = \rho_1^{(n)} v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} \left(1 - \mu^{(n)} - \sum_{k' \neq k} v_{k'} \mathbf{b}_{k'}^{\mathrm{T}} \mathbf{p}^{(n+1)}\right) + \rho_{2,k}^{(n)} \left(\left(\mathbf{z}^{(n)}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta\right) \left(1 - \gamma_k^{(n)}\right), \tag{57}$$

$$w_{k,2}^{(n+1)} = \rho_1^{(n)} \left(v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)}\right)^2 + \rho_{2,k}^{(n)} \left(\left(\mathbf{z}^{(n)}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta\right)^2. \tag{58}$$

- *Step 3: Update the primal variables* $\mathbf{z}$. Isolating all the terms involving $\mathbf{z}$, the optimization is an unconstrained convex problem

$$\min_{\mathbf{z}} \frac{\rho_0^{(n)}}{2} \left\| \mathbf{p}^{(n+1)} - \mathbf{z} + \mathbf{e}^{(n)} \right\|^2 + \sum_{k=1}^{K} \frac{\rho_{2,k}^{(n)}}{2} \left| w_k^{(n+1)} \mathbf{z}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1 + \gamma_k^{(n)} \right|^2. \tag{59}$$

Solving (59) yields the $(n+1)$th estimate of $\mathbf{z}$ as

$$\mathbf{z}^{(n+1)} = \left(\mathbf{Z}_1^{(n+1)}\right)^{-1} \mathbf{z}_2^{(n+1)}, \tag{60}$$

where

$$\mathbf{Z}_1^{(n+1)} = \rho_0^{(n)} \mathbf{I}_M + \sum_{k=1}^{K} \rho_{2,k}^{(n)} \left(w_k^{(n+1)} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta\right) \times \left(w_k^{(n+1)} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta\right)^{\mathrm{T}}, \tag{61}$$

$$\mathbf{z}_2^{(n+1)} = \rho_0^{(n)} \left(\mathbf{p}^{(n+1)} + \mathbf{e}^{(n)}\right) + \sum_{k=1}^{K} \rho_{2,k}^{(n)} \left(w_k^{(n+1)} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta\right) \left(1 - \gamma_k^{(n)}\right). \tag{62}$$

- *Step 4: Update the dual variables* $\mathbf{e}$, $\mu$ *and* $\boldsymbol{\gamma}$. Maximizing the Lagrangian (48) with respect to the dual variables yields

$$\mathbf{e}^{(n+1)} = \mathbf{e}^{(n)} + \mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)}, \tag{63}$$

$$\mu^{(n+1)} = \mu^{(n)} + \sum_{k=1}^{K} w_k^{(n+1)} v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} - 1, \tag{64}$$

$$\gamma_k^{(n+1)} = \gamma_k^{(n)} + w_k^{(n+1)} \left(\mathbf{z}^{(n+1)}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1, \quad 1 \le k \le K. \tag{65}$$

- *Step 5: Update the penalty parameters* $\rho_0$, $\rho_1$ *and* $\boldsymbol{\rho}_2$. The penalty parameters are updated at the end of each iteration for the first a few iterations to speed up the convergence. At the $(n+1)$th iteration, associated with the three penalty parameters of $\rho_0^{(n)}$, $\rho_1^{(n)}$ and $\boldsymbol{\rho}_2^{(n)}$, we have three primal residuals

$$r_0^{(n+1)} = \left\| \mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)} \right\|, \tag{66}$$

$$r_1^{(n+1)} = \left| \sum_{k=1}^{K} w_k^{(n+1)} v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} - 1 \right|, \tag{67}$$

$$r_{2,k}^{(n+1)} = \left| w_k \left(\mathbf{z}^{(n+1)}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1 \right|, \quad 1 \le k \le K, \tag{68}$$

as well as three respective dual residuals

$$s_0^{(n+1)} = \left\| \rho_0^{(n)} \left(\mathbf{z}^{(n+1)} - \mathbf{z}^{(n)}\right) \right\|, \tag{69}$$

$$s_1^{(n+1)} = \left\| \mathbf{s}_{1a}^{(n+1)} \right\|, \tag{70}$$

$$s_{2,k}^{(n+1)} = \sqrt{\left(s_{2a,k}^{(n+1)}\right)^2 + \left\| \mathbf{s}_{2b,k}^{(n+1)} \right\|}, \; 1 \le k \le K, \tag{71}$$

where

$$\mathbf{s}_{1a}^{(n+1)} = \mu^{(n+1)} \rho_1^{(n)} \left(\sum_{k=1}^{K} \left(w_k^{(n)} - w_k^{(n+1)}\right) v_k \mathbf{b}_k\right) + \rho_1^{(n)} \left(\sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k\right) \times \left(\sum_{k=1}^{K} \left(w_k^{(n)} - w_k^{(n+1)}\right) v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)}\right), \tag{72}$$

$$s_{2a,k}^{(n+1)} = \rho_{2,k}^{(n)} \left(\mathbf{z}^{(n)}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta \times \left(w_k^{(n+1)} \left(\mathbf{z}^{(n)} - \mathbf{z}^{(n+1)}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1\right) + \gamma_k^{(n+1)} \rho_{2,k}^{(n)} \left(\left(\mathbf{z}^{(n)} - \mathbf{z}^{(n+1)}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta\right), \tag{73}$$

$$\mathbf{s}_{2b,k}^{(n+1)} = \rho_{2,k}^{(n)} w_k^{(n)} \eta \mathbf{A}_k^{\mathrm{T}} \mathbf{z}^{(n)} \times \left(\left(w_k^{(n)} \left(\mathbf{z}^{(n)}\right)^{\mathrm{T}} - w_k^{(n+1)} \left(\mathbf{z}^{(n+1)}\right)^{\mathrm{T}}\right) \times \mathbf{A}_k \mathbf{p}^{(n+1)} \eta\right) + \gamma_k^{(n+1)} \rho_{2,k}^{(n)} \eta \mathbf{A}_k^{\mathrm{T}} \times \left(w_k^{(n)} \mathbf{z}^{(n)} - w_k^{(n+1)} \mathbf{z}^{(n+1)}\right). \tag{74}$$

The exact definitions of the dual residuals can be found in Appendix A.

The penalty parameter $\rho_0$ is updated as follows

$$\rho_0^{(n+1)} = \begin{cases} \tau \rho_0^{(n)}, & \text{if } r_0^{(n+1)} \ge \varepsilon s_0^{(n+1)}, \\ \frac{1}{\tau} \rho_0^{(n)}, & \text{if } s_0^{(n+1)} \ge \varepsilon r_0^{(n+1)}, \\ \rho_0^{(n)}, & \text{otherwise}, \end{cases} \tag{75}$$

where the scalars $\tau > 1$ and $\varepsilon > 1$ with typical values of $\tau = 2$ and $\varepsilon = 10$. The idea behind this penalty parameter update is to balance the primal and dual residual norms $r_0^{(n+1)}$ and $s_0^{(n+1)}$, i.e., to keep $\frac{r_0^{(n+1)}}{s_0^{(n+1)}} \le \varepsilon$ and $\frac{s_0^{(n+1)}}{r_0^{(n+1)}} \le \varepsilon$, as they both converge to zero [18], [25]. The related dual variables are rescaled to remove the impact of changing $\rho_0$ according to

$$\mathbf{e}^{(n+1)} = \frac{\rho_0^{(n)}}{\rho_0^{(n+1)}} \mathbf{e}^{(n)}. \tag{76}$$

Similarly, $\rho_1$ is updated according to

$$\rho_1^{(n+1)} = \begin{cases} \tau \rho_1^{(n)}, & \text{if } r_1^{(n+1)} \geq \varepsilon s_1^{(n+1)}, \\ \frac{1}{\tau} \rho_1^{(n)}, & \text{if } s_1^{(n+1)} \geq \varepsilon r_1^{(n+1)}, \\ \rho_1^{(n)}, & \text{otherwise.} \end{cases} \tag{77}$$

The related dual variable is then scaled according to

$$\mu^{(n+1)} = \frac{\rho_1^{(n)}}{\rho_1^{(n+1)}} \mu^{(n)}. \tag{78}$$

Likewise, $\rho_{2,k}$ for $1 \leq k \leq K$ are updated according to

$$\rho_{2,k}^{(n+1)} = \begin{cases} \tau \rho_{2,k}^{(n)}, & \text{if } r_{2,k}^{(n+1)} \geq \varepsilon s_{2,k}^{(n+1)}, \\ \frac{1}{\tau} \rho_{2,k}^{(n)}, & \text{if } s_{2,k}^{(n+1)} \geq \varepsilon r_{2,k}^{(n+1)}, \\ \rho_{2,k}^{(n)}, & \text{otherwise,} \end{cases} \tag{79}$$

and the corresponding dual variables are rescaled as

$$\gamma_k^{(n+1)} = \frac{\rho_{2,k}^{(n)}}{\rho_{2,k}^{(n+1)}} \gamma_k^{(n)}, \; 1 \leq k \leq K. \tag{80}$$

- *Termination of the iterative procedure*. The iterative procedure is terminated when $\left\|\mathbf{p}^{(n+1)} - \mathbf{p}^{(n)}\right\|$ becomes smaller than a predefined small positive value or the preset maximum number of iterations is reached. Otherwise, set $n = n+1$ and go to *Step 1*.

*Remark 2*: The ADMM combines the advantages of the dual ascent and the augmented Lagrangian method. The dual ascent approach deals with the complicated constraints, while the augmented Lagrangian method is capable of enhancing the convergence rate and the robustness of the algorithm.

*Remark 3*: We deal with the optimization problem (24), and in every iteration of our OCD and ADMM methods, we have a closed-form update value. By contrast, Garcia *et al.* [13] deal with the optimization problem (25), and in every iteration, an inner iterative loop is required for computing an updated value by the algorithm of [13].

### B. Single-Target Case

The target index $k$ can be dropped and then the optimization is simplified to the problem $\mathbb{P}5$

$$\mathbb{P}5: \begin{array}{ll} \min\limits_{\mathbf{p}} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & \frac{\mathbf{b}^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p}} \leq \eta, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \tag{81}$$

In the single-target case, the optimization (25) is identical to the problem $\mathbb{P}5$. Similar to the multi-target case, the problem $\mathbb{P}5$ is equivalent to the optimization problem $\mathbb{P}6$:

$$\mathbb{P}6: \begin{array}{ll} \min\limits_{\mathbf{p},w} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & w\mathbf{b}^{\mathrm{T}}\mathbf{p} - 1 = 0, \\ & w\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - 1 = 0, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \tag{82}$$

This problem is nonconvex due to its equality constraint.

*1) OCD-based method:* The Lagrangian of (82) is

$$L(\mathbf{p}, w, \lambda, \mu) = \mathbf{1}^{\mathrm{T}}\mathbf{p} + \lambda\left(w\mathbf{b}^{\mathrm{T}}\mathbf{p} - 1\right) + \mu\left(w\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - 1\right), \tag{83}$$

where $\lambda$ and $\mu$ are the dual variables. The gradients of this Lagrangian are given by

$$\Delta\mathbf{p} = \nabla_{\mathbf{p}} L(\mathbf{p}, w, \lambda, \mu) = \mathbf{1} + \lambda(w\mathbf{b}) + \mu w\eta\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)\mathbf{p}, \tag{84}$$

$$\Delta\lambda = \nabla_{-\lambda} L(\mathbf{p}, w, \lambda, \mu) = -w\mathbf{b}^{\mathrm{T}}\mathbf{p} + 1, \tag{85}$$

$$\Delta w = \nabla_{w} L(\mathbf{p}, w, \lambda, \mu) = \lambda\mathbf{b}^{\mathrm{T}}\mathbf{p} + \mu\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p}, \tag{86}$$

$$\Delta\mu = \nabla_{-\mu} L(\mathbf{p}, w, \lambda, \mu) = -\eta w\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - 1, \tag{87}$$

Given $\lambda^{(0)}$, $\mu^{(0)}$ and

$$\mathbf{p}^{(0)} = \mathbf{p}_{equ} = \frac{1}{\eta}\frac{\mathbf{b}^{\mathrm{T}}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{A}\mathbf{1}}\mathbf{1}, \tag{88}$$

$\mathbf{p}$, $\lambda$, $w$, $\mu$ are updated in the following iterative procedure

$$p_m^{(n+1)} = \left[p_m^{(n)} - \kappa_1 \Delta p_m^{(n)}\right]_{p_{m_{\min}}}^{p_{m_{\max}}}, \; 1 \leq m \leq M, \tag{89}$$

$$\lambda^{(n+1)} = \lambda^{(n)} - \kappa_2 \Delta\lambda^{(n)}, \tag{90}$$

$$w^{(n+1)} = w^{(n)} - \kappa_3 \Delta w^{(n)}, \tag{91}$$

$$\mu^{(n+1)} = \mu^{(n)} - \kappa_4 \Delta\mu^{(n)}, \tag{92}$$

where again the step sizes are chosen according to (42). The iterative procedure is repeated until $\left\|\mathbf{p}^{(n+1)} - \mathbf{p}^{(n)}\right\|$ becomes smaller than a preset threshold.

*2) ADMM-based method:* Similar to the multi-target case, we reformulate the problem $\mathbb{P}6$ as

$$\begin{array}{ll} \min\limits_{\mathbf{p},\mathbf{z}} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & \eta\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{p} - \mathbf{b}^{\mathrm{T}}\mathbf{p} = 0, \\ & \mathbf{z} = \mathbf{p}, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \tag{93}$$

Then, by introducing an augmented Lagrangian, we have

$$\begin{array}{ll} \max\limits_{\mathbf{e},\mu}\min\limits_{\mathbf{p},\mathbf{z}} & \mathbf{1}^{\mathrm{T}}\mathbf{p} + \frac{\rho_0}{2}\|\mathbf{p} - \mathbf{z} + \mathbf{e}\|^2 \\ & + \frac{\rho_1}{2}\left\|\eta\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{p} - \mathbf{b}^{\mathrm{T}}\mathbf{p} + \mu\right\|^2, \\ \text{s.t.} & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \tag{94}$$

With the initialization of $\mathbf{p}^{(0)} = \mathbf{z}^{(0)} = \mathbf{p}_{equ}$, $\mathbf{e}^{(0)} = \mathbf{1}$, $\mu^{(0)} = 1$, and $\rho_0^{(0)}$ and $\rho_1^{(0)}$ set to a large positive number, each iteration involves the following steps.

TABLE I
COMPLEXITY PER ITERATION OF THE OCD-BASED ALGORITHM

| Operation | Flops per iteration |
|---|---|
| Update $\mathbf{p}$ | $3KM^2 + (3K+5)M + 3K$ |
| Update $\lambda$ | $2KM + 2K + 2,$ |
| Update $\mathbf{w}$ | $2KM^2 + 3KM + 5K,$ |
| Update $\boldsymbol{\mu}$ | $2KM^2 + KM + 4K$ |
| Total | $7KM^2 + (9K+5)M + 14K + 2$ |

TABLE II
COMPLEXITY PER ITERATION OF THE ADMM-BASED ALGORITHM

| Operation | Flops per iteration |
|---|---|
| Update $\mathbf{p}$ | $M^3 + (5K+7)M^2 + (4K+8)M + 3K + 5$ |
| Update $\mathbf{w}$ | $4KM^2 + (2K^2+4K)M + K^2 + 14K$ |
| Update $\mathbf{z}$ | $M^3 + (7K+2)M^2 + (2K+3)M + 4K$ |
| Update $\mathbf{e}$ | $2M$ |
| Update $\mu$ | $2KM + 2K + 1$ |
| Update $\boldsymbol{\gamma}$ | $2KM^2 + KM + 3K$ |
| Total | $2M^3 + (18K+9)M^2 + \left(2K^2 + 13K + 13\right)M$ |
| | $+K^2 + 26K + 6$ |

- *Step 1: Update* $\mathbf{p}$. Isolating all the terms involving $\mathbf{p}$, the optimization is a constrained convex problem, leading to

$$\begin{aligned}
&\bar{\mathbf{p}}^{(n+1)} \\
&= \left(\rho_0^{(n)}\mathbf{I}_M + \rho_1^{(n)}\left(\eta\mathbf{A}^{\mathrm{T}}\mathbf{z}^{(n)} - \mathbf{b}\right)\right. \\
&\quad \left.\times\left(\eta\left(\mathbf{z}^{(n)}\right)^{\mathrm{T}}\mathbf{A} - \mathbf{b}^{\mathrm{T}}\right)\right)^{-1}\left(-\mathbf{1} + \rho_0^{(n)}\left(\mathbf{z}^{(n)} - \mathbf{e}^{(n)}\right)\right. \\
&\quad \left. - \rho_1^{(n)}\mu^{(n)}\left(\eta\mathbf{A}^{\mathrm{T}}\mathbf{z}^{(n)} - \mathbf{b}\right)\right), \qquad (95)
\end{aligned}$$

$$p_m^{(n+1)} = \left[\bar{p}_m^{(n+1)}\right]_{p_{m\,\min}}^{p_{m\,\max}}, \; 1 \le m \le M. \qquad (96)$$

- *Step 2: Update* $\mathbf{z}$. Isolating all the terms involving $\mathbf{z}$, the problem is an unconstrained convex problem, leading to

$$\begin{aligned}
\mathbf{z}^{(n+1)} = &\left(\rho_0^{(n)}\mathbf{I}_M + \rho_1^{(n)}\left(\eta\mathbf{A}\mathbf{p}^{(n+1)}\right)\left(\eta\mathbf{A}\mathbf{p}^{(n+1)}\right)^{\mathrm{T}}\right)^{-1} \\
&\times\left(\rho_0^{(n)}\left(\mathbf{p}^{(n+1)} + \mathbf{e}^{(n)}\right)\right. \\
&\left. + \rho_1^{(n)}\eta\mathbf{A}\mathbf{p}^{(n+1)}\left(\mathbf{b}^{\mathrm{T}}\mathbf{p}^{(n+1)} - \mu^{(n)}\right)\right). \qquad (97)
\end{aligned}$$

- *Step 3: Update* $\mathbf{e}$ *and* $\mu$. The dual variables are updated according to

$$\mu^{(n+1)} = \mu^{(n)} + \eta\left(\mathbf{z}^{(n+1)}\right)^{\mathrm{T}}\mathbf{A}\mathbf{p}^{(n+1)} - \mathbf{b}^{\mathrm{T}}\mathbf{p}^{(n+1)}, \quad (98)$$

$$\mathbf{e}^{(n+1)} = \mathbf{e}^{(n)} + \mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)}. \qquad (99)$$

- *Step 4: Update the* $\rho_0$ *and* $\rho_1$ *at the first a few iterations*. By defining the primal and dual residuals $r_0^{(n+1)}$ and $s_0^{(n+1)}$ as

$$r_0^{(n+1)} = \left\|\mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)}\right\|, \qquad (100)$$

$$s_0^{(n+1)} = \left\|\rho_0^{(n)}\left(\mathbf{z}^{(n)} - \mathbf{z}^{(n+1)}\right)\right\|, \qquad (101)$$

the updated $\rho_0^{(n+1)}$ is given by (75), and the dual variable $\mathbf{e}^{(n+1)}$ is rescaled according to (76). Similarly, define the primal and dual residuals $r_1^{(n+1)}$ and $s_1^{(n+1)}$ as

$$r_1^{(n+1)} = \left|\eta\left(\mathbf{z}^{(n+1)}\right)^{\mathrm{T}}\mathbf{A}\mathbf{p}^{(n+1)} - \mathbf{b}^{\mathrm{T}}\mathbf{p}^{(n+1)}\right|, \qquad (102)$$

$$\begin{aligned}
s_1^{(n+1)} = &\left\|\mu^{(n+1)}\rho_1^{(n)}\eta\mathbf{A}^{\mathrm{T}}\left(\mathbf{z}^{(n)} - \mathbf{z}^{(n+1)}\right) + \rho_1^{(n)}\eta\right. \\
&\left.\times\left(\eta\mathbf{A}^{\mathrm{T}}\mathbf{z}^{(n)} - \mathbf{b}\right)\left(\mathbf{z}^{(n)} - \mathbf{z}^{(n+1)}\right)^{\mathrm{T}}\mathbf{A}\mathbf{p}^{(n+1)}\right\|. \qquad (103)
\end{aligned}$$

The updated $\rho_1^{(n+1)}$ is given by (77), and the rescaled dual variable $\mu^{(n+1)}$ is given by (78).

*3) A closed-form approximate solution:* An equivalent Lagrangian associated with the problem $\mathbb{P}5$ is $L(\mathbf{p},\lambda) = \mathbf{1}^{\mathrm{T}}\mathbf{p} + \lambda\left(\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - \mathbf{b}^{\mathrm{T}}\mathbf{p}\right)$, whose KKT conditions are

$$\mathbf{1} + \lambda\left(\eta\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)\mathbf{p} - \mathbf{b}\right) = \mathbf{0}, \qquad (104)$$

$$\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - \mathbf{b}^{\mathrm{T}}\mathbf{p} = \mathbf{0}. \qquad (105)$$

The authors of [12] obtained the closed-form optimal solution $\lambda^\star$ and $\mathbf{p}^\star$ by jointly solving the two equations (104) and (105). In particular, they calculated $\bar{\mathbf{p}}^\star$ from (104) as

$$\bar{\mathbf{p}}^\star = \frac{\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)^{-1}}{\eta}\left(\mathbf{b} - \frac{1}{\lambda^\star}\mathbf{1}\right), \qquad (106)$$

and then obtained $\mathbf{p}^\star$ by imposing the power constraints

$$p_m^\star = \left[\bar{p}_m^\star\right]_{p_{m\,\min}}^{p_{m\,\max}}, \; 1 \le m \le M. \qquad (107)$$

Unfortunately, this closed-form 'optimal' solution is generally invalid because in general $\mathbf{A} + \mathbf{A}^{\mathrm{T}}$ is not invertible.

*Lemma 2:* The rank of $\mathbf{A} + \mathbf{A}^{\mathrm{T}}$ is no more than 3.

*Proof:*

$$\begin{aligned}
\mathrm{rank}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right) \le\; &\mathrm{rank}\left(\mathbf{a}_{1,1}\left(\mathbf{a}_{2,2}\right)^{\mathrm{T}} - \mathbf{a}_{1,2}\left(\mathbf{a}_{2,1}\right)^{\mathrm{T}}\right. \\
&\left. + \mathbf{a}_{2,2}\left(\mathbf{a}_{1,1}\right)^{\mathrm{T}} - \mathbf{a}_{2,1}\left(\mathbf{a}_{1,2}\right)^{\mathrm{T}}\right) \\
\le\; &\mathrm{rank}\left(\mathbf{a}_{1,1}\left(\mathbf{a}_{2,2}\right)^{\mathrm{T}}\right) + \mathrm{rank}\left(\mathbf{a}_{1,2}\left(\mathbf{a}_{2,1}\right)^{\mathrm{T}}\right) \\
&+ \mathrm{rank}\left(\mathbf{a}_{2,2}\left(\mathbf{a}_{1,1}\right)^{\mathrm{T}}\right) \le 3.
\end{aligned}$$

The second inequality is due to the fact that $\mathbf{a}_{1,2} = \mathbf{a}_{2,1}$.

Clearly, for any system with more than 3 transmit radars, the solution of (106) is invalid, and the minimum eigenvalue $\xi_{\min}$ of $\mathbf{A} + \mathbf{A}^{\mathrm{T}}$ is negative. We propose an approximate closed-form solution by replacing the invalid $\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)^{-1}$ in (106) by the valid regularized form

$$\mathbf{B} = \left(\mathbf{A} + \mathbf{A}^{\mathrm{T}} + \left(|\xi_{\min}| + \epsilon\right)\mathbf{I}_M\right)^{-1}, \qquad (108)$$

where $\epsilon$ is a small positive number, such as, $\epsilon = 0.0001$. Thus the 'unconstrained' power allocation is given as

$$\bar{\mathbf{p}}^\divideontimes = \frac{\mathbf{B}}{\eta}\left(\mathbf{b} - \frac{1}{\lambda^\divideontimes}\mathbf{1}\right), \qquad (109)$$

TABLE III
COMPLEXITY PER ITERATION OF THE ALGORITHM GIVEN IN [13], WHERE $n_{\text{in}}$ IS THE AVERAGE NUMBER OF INNER ITERATIONS IN INNER OPTIMIZATION PROCEDURE

| Operation | Flops per inner iteration |
|---|---|
| Update the parameters of inner QCLP problem | $(5M^2+2M+1)/n_{\text{in}}$ |
| Solve the inner QCLP problem | $M^3+(4K+3)M^2+(6K+10)M-K$ |
| Total | $M^3+\left(4K+3+\frac{5}{n_{\text{in}}}\right)M^2+\left(6K+10+\frac{2}{n_{\text{in}}}\right)M-K$ |

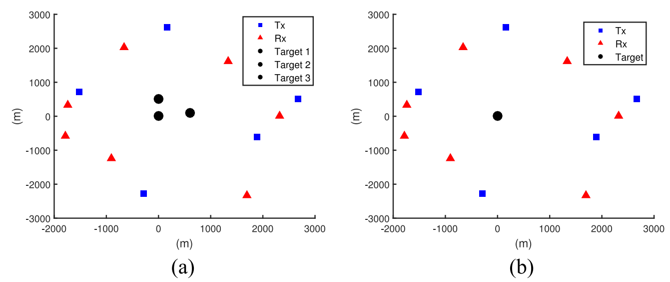



Fig. 2. Illustration of the MIMO radar system for: (a) three-target application, and (b) single-target application.

where $\lambda^{\divideontimes}$ is obtained by substituting $\bar{\mathbf{p}}^{\divideontimes}$ into (105) and taking the positive solution as

$$\lambda^{\divideontimes} = \frac{-b+\sqrt{b^2-4ac}}{2a}, \tag{110}$$

with

$$\begin{cases} a = \mathbf{b}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}\mathbf{B}\mathbf{b} - \mathbf{b}^{\mathrm{T}}\mathbf{B}\mathbf{b}, \\ b = -2\mathbf{1}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{B}\mathbf{b}^{\mathrm{T}} + 2\mathbf{b}^{\mathrm{T}}\mathbf{B}\mathbf{1}, \\ c = \mathbf{1}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}\mathbf{B}\mathbf{1} - \mathbf{1}^{\mathrm{T}}\mathbf{B}\mathbf{1}. \end{cases} \tag{111}$$

The solution $\mathbf{p}^{\divideontimes}$ is then obtained by projecting $\bar{\mathbf{p}}^{\divideontimes}$ onto the feasible region. This closed-form solution is inferior to the OCD-based and ADMM-based solutions in terms of its achievable performance, owing to its suboptimal nature.

## IV. CONVERGENCE AND COMPLEXITY ANALYSIS

Recall from Section II and III that our optimization problem $\mathbb{P}1$ of (24) is nonconvex, and both our ADMM and OCD algorithms are based on a Lagrangian function approach. It is widely acknowledged that the zero duality gap cannot be guaranteed for general nonconvex problems. However, Yu and Lui [24] proposed a theorem which guarantees the zero duality gap for the nonconvex problem that meets the 'time-sharing condition'. In Appendix B, we proved that our optimization problem $\mathbb{P}1$ satisfies the time-sharing condition of [24]. Hence, the strong duality holds for $\mathbb{P}1$. We are now ready to prove that both our two algorithms can converge to a local optimal point under some assumptions.

### A. Convergence of the Proposed Algorithms

*1) The ADMM-based algorithm:* We first point out again that since our problem is nonconvex, the ADMM-based algorithm can only guarantee to converge to a local optimal solution. The convergence of the ADMM method is proved for the convex optimization problem in [18], while Magnússon *et al.* [25] extended the convergence results to the nonconvex case. The convergence of our ADMM-based algorithm will be further illustrated in Section V using simulations.

*2) The OCD-based algorithm:* Again, since our optimization problem is nonconvex, the OCD-based algorithm can only find a locally optimal solution. Collect all the primal variables of the Lagrangian (32) together as $\mathbf{y} = \left[\mathbf{p}^{\mathrm{T}}\ \mathbf{w}^{\mathrm{T}}\right]^{\mathrm{T}}$ and denote the cost function and the constraints of $\mathbb{P}3$ respectively by

$$f(\mathbf{y}) = \mathbf{1}^{\mathrm{T}}\mathbf{p}, \tag{112}$$

$$g_0(\mathbf{y}) = \sum_{k=1}^{K} v_k w_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1, \tag{113}$$

$$g_k(\mathbf{y}) = w_k \eta \mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} - 1,\ 1 \le k \le K. \tag{114}$$

According to Theorem 2 in Section 8.2.3 and Lemma 5 in Section 2.1 of [26], to prove the convergence of the OCD algorithm, we have to prove that the second derivatives $\nabla^2 f(\mathbf{y})$ and $\nabla^2 g_k(\mathbf{y})$ for $0 \le k \le K$ satisfy the Lipschitz condition in a neighbourhood of the optimal primal point $\mathbf{y}^\star$. Note that

$$\nabla^2 f(\mathbf{y}) = \mathbf{0}, \tag{115}$$

$$\nabla^2 g_0(\mathbf{y}) = \begin{bmatrix} \mathbf{0} & v_1\mathbf{b}_1 \cdots v_K\mathbf{b}_K \\ v_1\mathbf{b}_1^{\mathrm{T}} & \\ \vdots & \mathbf{0} \\ v_K\mathbf{b}_K^{\mathrm{T}} & \end{bmatrix}, \tag{116}$$

$$\nabla^2 g_k(\mathbf{y}) = \eta \begin{bmatrix} w_k\left(\mathbf{A}_k+\mathbf{A}_k^{\mathrm{T}}\right) & \mathbf{0} & \left(\mathbf{A}_k+\mathbf{A}_k^{\mathrm{T}}\right)\mathbf{p} & \mathbf{0} \\ \mathbf{0} & & & \\ \left(\mathbf{A}_k+\mathbf{A}_k^{\mathrm{T}}\right)\mathbf{p}^{\mathrm{T}} & & \mathbf{0} & \\ \mathbf{0} & & & \end{bmatrix}, \quad 1 \le k \le K. \tag{117}$$

Since $\nabla^2 f(\mathbf{y})$ and $\nabla^2 g_0(\mathbf{y})$ are constants, they satisfy the required Lipschitz condition. For $\mathbf{p}_{\min} \le \mathbf{p} \le \mathbf{p}_{\max}$, all the elements in the matrix $\nabla^2 g_k(\mathbf{y})$, where $1 \le k \le K$, are finite. Therefore, it is easy to find a constant $\varsigma$ satisfying

$$\left\|\nabla^2 g_k(\mathbf{y}_1) - \nabla^2 g_k(\mathbf{y}_2)\right\| \le \varsigma \left\|\mathbf{y}_1 - \mathbf{y}_2\right\|. \tag{118}$$

Thus $\nabla^2 g_k(\mathbf{y})$ satisfies the required Lipschitz condition too.

According to [26], under the assumption that the Hessian matrix of the Lagrangian (32) with respect to $\mathbf{y}$ at the minimum primal point $\mathbf{y}^\star = \left(\mathbf{p}^\star, \mathbf{w}^\star\right)$ is positive definite, the Hessian matrix of the Lagrangian (32) with respect to the primal and dual variables is negative definite at the optimal point $\left(\mathbf{p}^\star, \mathbf{w}^\star, \lambda^\star, \boldsymbol{\mu}^\star\right)$. Then there exists a positive number $\overline{\kappa} = \min_i -\Re\left[\bar{\xi}_i\right]\left|\bar{\xi}_i\right|^{-2}$,

TABLE IV
SYSTEM PARAMETERS

| Parameters | | Values |
|---|---|---|
| Effective bandwidth $\beta_m$ | | 200 MHz |
| Transmit power upper bound $p_{m_{\max}}$ | | 300 watts |
| Transmit power lower bound $p_{m_{\min}}$ | | 1 watts |
| Transmit radars' positions | | $(2665, 508)$, $(165, 2617)$, $(-1520, 715)$, $(-287, -2270)$, $(1892, -615)$ |
| Receive radars' positions | | $(2320, 0)$, $(1338, 1617)$, $(-656, 2019)$, $(-1740, 332)$, $(-1791, -582)$, $(-900, -1238)$, $(1696, -2334)$ |
| Path loss $\kappa_{m,n}^{(k}$ | | $\kappa_{m,n}^{(k} = \frac{1}{10\left(R_{m,k}^{tx}\right)^2\left(R_{n,k}^{rx}\right)^2}$ |
| Three targets | Targets' positions $(x_k, y_k)$ | $(0, 0)$, $(0, 500)$, $(600, 100)$ |
| | RCS model for target 1, $\left\|\mathbf{h}^{(1}\right\|$ | $\begin{matrix} 0.75 & 0.4 & 0.45 & 0.55 & 0.3 & 0.2 & 0.25 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.1 & 0.05 & 0.01 & 0.12 & 0.09 & 0.2 & 0.19 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$ |
| | RCS model for target 2, $\left\|\mathbf{h}^{(2}\right\|$ | $\begin{matrix} 0.1 & 0.05 & 0.01 & 0.12 & 0.09 & 0.2 & 0.19 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.75 & 0.4 & 0.45 & 0.55 & 0.3 & 0.2 & 0.25 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$ |
| | RCS model for target 3, $\left\|\mathbf{h}^{(3}\right\|$ | $\begin{matrix} 0.1 & 0.05 & 0.01 & 0.12 & 0.09 & 0.2 & 0.19 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.75 & 0.4 & 0.45 & 0.55 & 0.3 & 0.2 & 0.25 \end{matrix}$ |
| Single target | Target's position $(x, y)$ | $(0, 0)$ |
| | RCS model for target, $\left\|\mathbf{h}\right\|$ | $\begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.01 & 0.05 & 0.01 & 0.022 & 0.092 & 0.092 & 0.092 \\ 0.45 & 0.35 & 0.48 & 0.32 & 0.49 & 0.49 & 0.49 \\ 0.22 & 0.55 & 0.55 & 0.48 & 0.57 & 0.57 & 0.57 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$ |



where $\bar{\xi}_i$ are the eigenvalues of the Hessian matrix of the Lagrangian (32) with respect to the primal and dual variables at $\left(\mathbf{p}^\star, \mathbf{w}^\star, \lambda^\star, \boldsymbol{\mu}^\star\right)$. Consequently, as long as the maximum of the four step sizes $\kappa_{\max} = \max\limits_{1 \le i \le 4} \kappa_i$ satisfies the condition of $\kappa_{\max} \le \overline{\kappa}$, our scheme (37)–(40) will converge to the locally optimal point $\left(\mathbf{p}^\star, \mathbf{w}^\star, \lambda^\star, \boldsymbol{\mu}^\star\right)$ when starting from a neighbourhood of $\left(\mathbf{p}^\star, \mathbf{w}^\star, \lambda^\star, \boldsymbol{\mu}^\star\right)$, according to [26]. In practice, $\overline{\kappa}$ is unknown. It is advisable to choose sufficiently small step sizes $\kappa_i$, $1 \le i \le 4$, in order to ensure the convergence of the OCD scheme.

*Remark 4*: A positive-definite Hessian matrix of the Lagrangian (32) with respect to $\mathbf{y}$ at $\mathbf{y}^\star$ is a sufficient condition for the convergence of the OCD scheme. If this Hessian matrix is semi-positive definite, we cannot prove the convergence of the OCD scheme based on the result of [26]. By adopting an exponentially decaying step size $\kappa_{\max} \propto e^{-\alpha n}$, we ensure that our OCD algorithm works well in any situation.

### B. *Complexity of Proposed Algorithms and Algorithm of [13]*

The complexity of our OCD and ADMM algorithms are summarized in Tables I and II, respectively. For the ADMM-based algorithm, since the penalty parameters are only updated in the first few iterations, the complexity associated with this part of operation is omitted. Additionally, we assume that Gauss-Jordan elimination is used for matrix inversion and, therefore, the number of flops required by inverting an $M \times M$ matrix is $M^3 + M^2 + M$. For the OCD-based algorithm, the complexity of computing the four step sizes is negligible and therefore it is also omitted. Clearly, the complexity of the ADMM-based algorithm is on the order of $M^3$ per iteration, which is denoted by $\mathsf{O}\left(M^3\right)$, while the complexity of the OCD-based algorithm is on the order of $\mathsf{O}\left(M^2\right)$ per iteration. It will be shown by our simulation results that the convergence speed of the ADMM algorithm is at least one order of magnitude faster than that of the OCD algorithm. Therefore, despite its higher per-iteration complexity, the ADMM algorithm actually imposes a lower total complexity, compared to the OCD algorithm.

The benchmark scheme of [13] invokes two iterative loops for solving the optimization problem (25). Specifically, at each outer iteration, the parameters of the inner quadratic constrained linear programming (QCLP) problem are updated, and the QCLP problem is then solved iteratively in the inner iterative loop. We assume that the interior-point method is used for solving this inner QCLP, which requires $n_{\text{in}}$ iterations on average. Based on the above discussions, the complexity of the algorithm of [13] is summarized in Table III, where it is seen that the complexity per inner iteration is on the order of $\mathsf{O}\left(M^3\right)$. Thus the complexity of our ADMM-based algorithm is only marginally higher than that of the algorithm in [13], because they are both on the order of $\mathsf{O}\left(M^3\right)$ per iteration. The algorithm of [13] requires a total of $n_{\text{ou}} n_{\text{in}}$ iterations to converge, where $n_{\text{ou}}$ is the number of iterations for the outer iterative loop. As it will be shown in the simulation results, the number of iterations required for the ADMM-based algorithm to converge is very close to the total number of iterations $n_{\text{ou}} n_{\text{in}}$ required by the algorithm of [13].

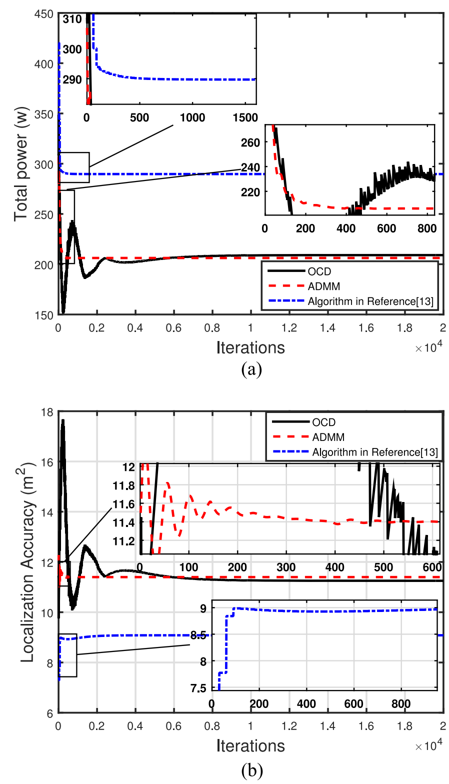



Fig. 3. Convergence performance of three algorithms, in terms of (a) total power consumption, and (b) aggregate localization accuracy, for the three-target case with $v_1 = 1$, $v_2 = 2$ and $v_3 = 1$.

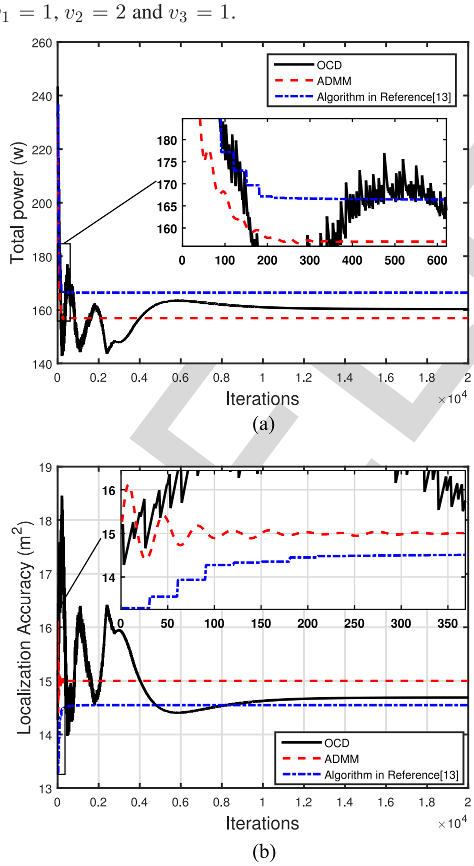



Fig. 4. Convergence performance of three algorithms, in terms of (a) total power consumption, and (b) aggregate localization accuracy, for the three-target case with $v_1 = v_2 = v_3 = 1$.

In this sense, both algorithms require a similar total complexity for solving their associated optimization problems. Although our OCD-based algorithm enjoys a much lower complexity per iteration than the algorithm of [13], it imposes a higher total complexity.

## V. Simulation Results

Let us now evaluate the performance of the proposed algorithms using a MIMO radar system having $M = 5$ transmit radars and $N = 7$ receive radars. The algorithm of [13] is used as the benchmark. Fig. 2 depicts both the triple-target and single-target cases considered. The system parameters of both the triple-target and single-target cases are listed in Table IV. The localization accuracy threshold $\eta$ is set to 15 m$^2$ for the triple-target case and 10 m$^2$ for the single-target case. The exponential decaying factor is empirically chosen to be $\alpha = 0.0005$ for the four step sizes of the OCD algorithm.

### A. Triple-Target Case

We consider the two sets of the importance weightings for the three targets given by: i) $v_1 = 1$, $v_2 = 2$ and $v_3 = 1$, and ii) $v_1 = v_2 = v_3 = 1$. For the sake of a fair comparison to the algorithm of [13], the effects of these weightings have to be taken into consideration, and the target estimation error thresholds for the three constraints of the optimization problem (25) are suitably scaled as

$$\frac{\mathbf{b}_1^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_1\mathbf{p}} \le \bar{\eta}_1, \ \frac{\mathbf{b}_2^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_2\mathbf{p}} \le \bar{\eta}_2, \ \frac{\mathbf{b}_3^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_3\mathbf{p}} \le \bar{\eta}_3,$$

with $\bar{\eta}_1 = \frac{1}{3v_1}\eta$, $\bar{\eta}_2 = \frac{1}{3v_2}\eta$ and $\bar{\eta}_3 = \frac{1}{3v_3}\eta$. For our ADMM algorithm, the initial values of the dual variables are set to $\mathbf{e}^{(0)} = [1\,1\,1\,1\,1]^{\mathrm{T}}$, $\mu^{(0)} = 1$ and $\gamma_k^{(0)} = 1$ for $1 \le k \le 3$, while all the initial penalty parameters are set to 500. For our OCD algorithm, the initial values of the dual variables are set to $\lambda^{(0)} = 1$ and $\mu_k^{(0)} = 1$ for $1 \le k \le 3$. Additionally, the four constants in the four step sizes of the OCD algorithm are set to $c_1 = 0.3$, $c_2 = 1.0$, $c_3 = 1.5$ and $c_4 = 1.1$ for the senario i), while they are set to $c_1 = 0.3$, $c_2 = 0.9$, $c_3 = 1.5$ and $c_4 = 1.1$ for the senario ii). These parameters were found empirically to be appropriate for the corresponding application scenarios. For the algorithm of [13], we use the CVX software to solve its inner QCLP problem. In our simulations, we observe that the CVX converges within 25 to 35 iterations. Therefore, we will assume that the average number of inner iterations for the algorithm of [13] is $n_{\mathrm{in}} = 30$. Q1

Fig. 3 compares the total power allocations $\mathbf{p}$ and the aggregate localization accuracy results of $\sum_{k=1}^{3}\frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}}$ obtained by the three algorithms for the senario i), while Fig. 4 depicts the results for the senario ii). It can be seen that the number of iterations required by the ADMM-based algorithm to converge is similar to the total number of iterations $n_{\mathrm{ou}}n_{\mathrm{in}}$ required by the algorithm of [13], while the convergence speed of the OCD-based algorithm is considerably slower than that of the other two algorithms. As expected, our algorithms outperform the algorithm of [13] in terms of its total power consumption, albeit at the expense of some degradation in localization accuracy.

TABLE V
PERFORMANCE COMPARISON OF THREE ALGORITHMS FOR THE THREE-TARGET CASE

| | ii) $v_1 = v_2 = v_3 = 1$ | | | i) $v_1 = 1,\ v_2 = 2,\ v_3 = 1$ | | |
|---|---|---|---|---|---|---|
| | ADMM | OCD | [13] | ADMM | OCD | [13] |
| Radar 1: Power (watts) | 1 | 1 | 1 | 1 | 1 | 1 |
| Radar 2: Power (watts) | 95.8 | 93.3 | 102 | 119.6 | 117.9 | 75.8 |
| Radar 3: Power (watts) | 58.2 | 64.0 | 40.3 | 83.5 | 88.1 | 170.4 |
| Radar 4: Power (watts) | 1 | 1 | 1 | 1 | 1 | 1 |
| Radar 5: Power (watts) | 1 | 1 | 22.2 | 1 | 1 | 41.5 |
| Total Power (watts) | 157 | 160.3 | 166.5 | 206.1 | 209.0 | 289.7 |
| Target 1: Localization Accuracy ($m^2$) | 5.4 | 5.3 | 5 | 4.1 | 4.1 | 3.1 |
| Target 2: Localization Accuracy ($m^2$) | 4.8 | 4.7 | 4.5 | 3.6 | 3.5 | 2.5 |
| Target 3: Localization Accuracy ($m^2$) | 4.8 | 4.7 | 5 | 3.7 | 3.6 | 3.5 |
| Aggregate Localization Accuracy ($m^2$) | 15 | 14.7 | 14.5 | 11.4 | 11.3 | 9.1 |
| Total Power Saving | 5.7% | 3.7% | - | 28.9% | 27.9% | - |
| Degradation in Aggregate Localization Accuracy | 3.4% | 1.4% | - | 25.3% | 27.9% | - |
| Average Total Power Saving | 10.0% | 10.5% | - | 20.0 % | 25.6 % | - |
| Average Degradation in Aggregate Localization Accuracy | 8.6% | 8.9% | - | 27.2% | 30.0% | - |

The average results are obtained over 1000 random simulation experiments.

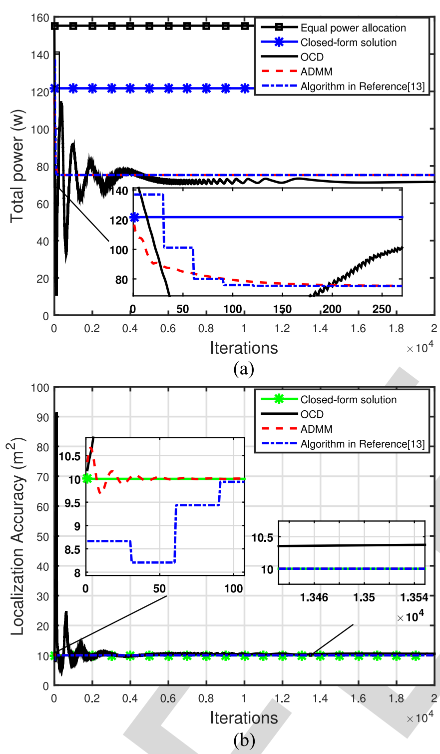



Fig. 5. Convergence performance of three algorithms, in terms of (a) total power consumption, and (b) aggregate localization accuracy, in comparison with the EPA and the closed-form solution, for the single-target case.

Table V details how our algorithms trade the localization accuracy against the transmit power, in comparison to the algorithm of [13]. Specifically, for the senario of i), our ADMM algorithm achieves 28.9% power saving at the cost of 25.3% degradation in aggregate localization accuracy, while our OCD algorithm trades 27.9% power saving against 27.9% degradation in localization accuracy. For the equal weighting senario of ii), the savings in power achieved by our two algorithms are considerably smaller but the losses in localization accuracy are also significantly smaller, compared with the senario i). To obtain statistically relevant comparison, we carry out 1000 simulations by randomly locating all the transmit radars and receive radars at the radius $R = 3000(0.5 + \varepsilon_x)$ m with the angular rotations of $\theta = 2\pi\varepsilon_y$, where $\varepsilon_x$ and $\varepsilon_y$ are uniformly distributed in $[0,\ 1.0]$. The average power saving and degradation in localization accuracy over the 1000 random experiments are listed in the last two rows of Table V.

## *B. Single-Target Case*

The four constants in the four step sizes of the OCD algorithm are set to $c_1 = c_2 = 1.0$ and $c_3 = c_4 = 0.3$, which is empirically found to be appropriate for this application senario. Fig. 5 characterizes the performance of our ADMM-based and OCD-based algorithms as well as the algorithm of [13]. As expected, all the three algorithms attain the same performance, both in terms of total power allocated and localization accuracy, since the underlying optimization problems are identical in the single-target case. In terms of convergence speed, our ADMM-based algorithm outperforms the algorithm of [13], while the OCD-based algorithm is considerably slower than the algorithm of [13]. In Fig. 5 (a), we also characterize the equal-power allocation (EPA) scheme and the closed-form solution of SubSection III-B3. It can be seen that our closed-form solution performs significantly better than the EPA scheme, but it is inferior to the other three iterative algorithms because the suboptimal nature of this closed-form solution.

## VI. CONCLUSION

The target localization problem of distributed MIMO radar systems has been investigated, which minimizes the power of the transmit radars, while meeting a required localization accuracy. We have proposed the OCD-based and ADMM-based iterative algorithms to solve this nonconvex optimization problem. Both the algorithms are capable of converging to a local optimum. The OCD algorithm imposes a much lower computational complexity per iteration, while the ADMM algorithm achieves a much faster convergence. For the multi-target senario, we have shown how our proposed approach trades the power saving with some degradation in localization accuracy, compared with that of state-of-the-art scheme [13]. We have also demonstrated that our ADMM-based algorithm and the existing state-of-the-art scheme have similar computational complexity and convergence speed. For the single-target senario, we have confirmed that our algorithms and the benchmark attain the same performance in terms of both power consumption and localization accuracy, because the underlying optimization problems become identical.

## Appendix

### A. Derivation of Updating Formulae for Penalty Parameters

The optimal solution to the $\mathbb{P}4$ of (45) should be primal and dual feasible, that is,

$$\mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)} = \mathbf{0}, \tag{119}$$

$$\sum_{k=1}^{K} w_k^{(n+1)} v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} - 1 = 0, \tag{120}$$

$$w_k \left(\mathbf{z}^{(n+1)}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1 = 0,\ 1 \le k \le K, \tag{121}$$

$$\frac{\partial L'\left(\mathbf{p}, \mathbf{z}^{(n+1)}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{p}} = \mathbf{0}, \tag{122}$$

$$\frac{\partial L'\left(\mathbf{p}^{(n+1)}, \mathbf{z}^{(n+1)}, \mathbf{w}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{w}} = \mathbf{0}, \tag{123}$$

$$\frac{\partial L'\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{z}} = \mathbf{0}, \tag{124}$$

where $L'\left(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_0, d_1, \mathbf{d}_2\right)$ is the Lagrangian of (45), which can be separated into three parts

$$L'\left(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_0, d_1, \mathbf{d}_2\right) = \underbrace{\mathbf{1}^{\mathrm{T}}\mathbf{p} + \mathbf{d}_0^{\mathrm{T}}(\mathbf{p} - \mathbf{z})}_{L_0'(\mathbf{p},\mathbf{z},\mathbf{d}_0)} + \underbrace{d_1 \left(\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p} - 1\right)}_{L_1'(\mathbf{p},\mathbf{w},d_1)} + \underbrace{\sum_{k=1}^{K} d_{2,k} \left(w_k \mathbf{z}^{\mathrm{T}} \mathbf{A}_k \mathbf{p} \eta - 1\right)}_{L_2'(\mathbf{p},\mathbf{w},\mathbf{z},\mathbf{d}_2)}. \tag{125}$$

However, the ADMM-based algorithm uses the augmented Lagrangian of

$$\begin{aligned} L(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_0, d_1, \mathbf{d}_2) = {} & \underbrace{\mathbf{1}^{\mathrm{T}}\mathbf{p} + \frac{\rho_0}{2}\|\mathbf{p} - \mathbf{z}\|^2 + \mathbf{d}_0^{\mathrm{T}}(\mathbf{p} - \mathbf{z})}_{L_0(\mathbf{p},\mathbf{z},\mathbf{d}_0)} \\ & + \underbrace{\frac{\rho_1}{2}\left|\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p} - 1\right|^2 + d_1 \left(\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p} - 1\right)}_{L_1(\mathbf{p},\mathbf{w},d_1)} \\ & + \underbrace{\sum_{k=1}^{K} \frac{\rho_{2,k}}{2}\left|w_k \mathbf{z}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}\eta - 1\right|^2 + \sum_{k=1}^{K} d_{2,k}\left(w_k \mathbf{z}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}\eta - 1\right)}_{L_2(\mathbf{p},\mathbf{w},\mathbf{z},\mathbf{d}_2)}, \end{aligned} \tag{126}$$

which can be divided into three parts, and all the primal and dual variables are updated one by one. Thus, in every iteration, there exist primal and dual residuals.

Specifically, in the $(n+1)$th iteration, the primal residuals are given by $r_0^{(n+1)}$ of (65), $r_1^{(n+1)}$ of (66), and $r_{2,k}^{(n+1)}$ for $1 \le k \le K$ of (67), while the dual residuals are defined via

$$dr = \sqrt{\|\mathbf{dr}_0\|^2 + \|\mathbf{dr}_1\|^2 + \|\mathbf{dr}_2\|^2}, \tag{127}$$

with

$$\mathbf{dr}_0 = \frac{\partial L\left(\mathbf{p}, \mathbf{z}^{(n)}, \mathbf{w}^{(n)}, \mathbf{d}_0^{(n)}, d_1^{(n)}, \mathbf{d}_2^{(n)}\right)}{\partial \mathbf{p}} - \frac{\partial L'\left(\mathbf{p}, \mathbf{z}^{(n+1)}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{p}}, \tag{128}$$

$$\mathbf{dr}_1 = \frac{\partial L\left(\mathbf{p}^{(n+1)}, \mathbf{z}^{(n)}, \mathbf{w}, \mathbf{d}_0^{(n)}, d_1^{(n)}, \mathbf{d}_2^{(n)}\right)}{\partial \mathbf{w}} - \frac{\partial L'\left(\mathbf{p}^{(n+1)}, \mathbf{z}^{(n+1)}, \mathbf{w}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{w}}, \tag{129}$$

$$\mathbf{dr}_2 = \frac{\partial L\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n)}, d_1^{(n)}, \mathbf{d}_2^{(n)}\right)}{\partial \mathbf{z}} - \frac{\partial L'\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{z}}. \tag{130}$$

It can be seen that the primal residuals $r_0^{(n+1)}$, $r_1^{(n+1)}$ and $r_{2,k}^{(n+)}$ for $1 \le k \le K$ are related to $L_0(\mathbf{p}, \mathbf{z}, \mathbf{d}_0)$, $L_1(\mathbf{p}, \mathbf{w}, d_1)$ and $L_2(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_2)$, respectively. Therefore, we will similarly 'separate' the dual residuals into $s_0^{(n+1)}$, $s_1^{(n+1)}$ and $s_{2,k}^{(n+1)}$ for $1 \le k \le K$, corresponding to $L_0(\mathbf{p}, \mathbf{z}, \mathbf{d}_0)$, $L_1(\mathbf{p}, \mathbf{w}, d_1)$ and $L_2(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_2)$, respectively.

In order to analyze the updating formula (75) for the penalty parameter $\rho_0$, we have to calculate $s_0^{(n+1)}$ as follows

$$s_0^{(n+1)} = \left( \left\| \frac{\partial L_0\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{d}_0^{(n)}\right)}{\partial \mathbf{z}} - \frac{\partial L_0'\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{d}_0^{(n+1)}\right)}{\partial \mathbf{z}} \right\|^2 + \left\| \frac{\partial L_0\left(\mathbf{p}, \mathbf{z}^{(n)}, \mathbf{d}_0^{(n)}\right)}{\partial \mathbf{p}} - \frac{\partial L_0'\left(\mathbf{p}, \mathbf{z}^{(n+1)}, \mathbf{d}_0^{(n+1)}\right)}{\partial \mathbf{p}} \right\|^2 \right)^{\frac{1}{2}}. \tag{131}$$

By evaluating the required four partial derivatives and plugging them into (131), we arrive at the dual residual $s_0^{(n+1)}$ of (68). Note that a large value for $\rho_0$ adds a large penalty on the violation of primal feasibility and, therefore, a large $\rho_0$ reduces the primal residual $r_0^{(n+1)}$. On the other hand, from the expression (68), it is seen that a small $\rho_0$ reduces the dual residual $s_0^{(n+1)}$. Thus, in order to balance the primal and dual residuals $r_0^{(n+1)}$ and $s_0^{(n+1)}$, the penalty parameter $\rho_0$ is updated according to (75), which is beneficial to convergence.

Similarly, it can be shown that the dual residual $s_1^{(n+1)}$ related to $L_1(\mathbf{p}, \mathbf{w}, d_1)$ is given by (69) and (71), while the dual residuals $s_{2,k}^{(n+1)}$ for $1 \leq k \leq K$ related to $L_2(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_2)$ are specified by (70), (72) and (73). Following the same logic of balancing the primal and dual residuals, the updating formulae for the penalty parameters $\rho_1$ and $\rho_{2,k}$ are specified by (76) and (78), respectively.

### B. *Proof of the Time-Sharing Condition for Problem* $\mathbb{P}1$

According to [24], the time-sharing condition for the optimization problem $\mathbb{P}1$ of (24) is as follows. *Time-sharing condition*: Let $\mathbf{p}_1$ and $\mathbf{p}_2$ be the optimal solutions of $\mathbb{P}1$ in conjunction with $\eta = \eta_1$ and $\eta = \eta_2$, respectively. $\mathbb{P}1$ is said to satisfy the time-sharing condition if for any $\eta_1$ and $\eta_2$ and for any $0 \leq \xi \leq 1$, there always exists a feasible solution $\mathbf{p}_3$ so that $\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}} \mathbf{A}_k \mathbf{p}_3} \leq \xi\eta_1 + (1-\xi)\eta_2$ and $\mathbf{1}^{\mathrm{T}}\mathbf{p}_3 \geq \xi\mathbf{1}^{\mathrm{T}}\mathbf{p}_1 + (1-\xi)\mathbf{1}^{\mathrm{T}}\mathbf{p}_2$.

According to Lemma 1, if we set $\mathbf{p}_3 = \mathbf{p}_{\max}$, then

$$\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}} \mathbf{A}_k \mathbf{p}_3} \leq \eta_1 \text{ and } \sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}} \mathbf{A}_k \mathbf{p}_3} \leq \eta_2.$$

Hence

$$\begin{aligned}\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}} \mathbf{A}_k \mathbf{p}_3} &= \xi \sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}} \mathbf{A}_k \mathbf{p}_3} \\ &\quad + (1-\xi) \sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}} \mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}} \mathbf{A}_k \mathbf{p}_3} \leq \xi\eta_1 + (1-\xi)\eta_2, \\ \mathbf{1}^{\mathrm{T}}\mathbf{p}_3 = \xi\mathbf{1}^{\mathrm{T}}\mathbf{p}_3 + (1-\xi)\mathbf{1}^{\mathrm{T}}\mathbf{p}_3 &\geq \xi\mathbf{1}^{\mathrm{T}}\mathbf{p}_1 + (1-\xi)\mathbf{1}^{\mathrm{T}}\mathbf{p}_2.\end{aligned}$$

Therefore, $\mathbb{P}1$ satisfies the time-sharing condition and the dual gap for our nonconvex problem is zero.

## REFERENCES


[1] J. Li and P. Stoica, *MIMO Radar Signal Processing*. Hoboken, NJ, USA: Wiley, 2009.

[2] Y. Yu, A. P. Petropulu, and H. V. Poor, "Measurement matrix design for compressive sensing–based MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5338–5352, Nov. 2011.

[3] S. Gogineni and A. Nehorai, "Target estimation using sparse modeling for distributed MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5315–5325, Nov. 2011.

[4] Y. Yao, A. P. Petrppulu, and H. V. Poor, "CSSF MIMO radar: Compressive-sensing and step-frequency based MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1490–1504, Apr. 2012.

[5] J. Xu, X.-Z. Dai, X.-G. Xia, L.-B. Wang, J. Yu, and Y.-N. Peng, "Optimization of multisite radar system with MIMO radars for target detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 4, pp. 2329–2343, Oct. 2011.

[6] E. Fishler, A. Haimovich, R. S. Blum, L. J. Cimini, D. Chizhik, and R. A. Valenzuela, "Spatial diversity in radars-models and detection performance," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 823–838, Mar. 2006.

[7] Q. He, N. H. Lehmann, R. S. Blum, and A. M. Haimovich, "MIMO-radar moving target detection in homogeneous clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 3, pp. 1290–1301, Jul. 2010.

[8] Q. He, R. S. Blum, H. Godrich, and A. M. Haimovich, "Cramer-Rao bound for target velocity estimation in MIMO radar with widely separated antennas," in *Proc. 42nd Annu. Conf. Inf. Sci. Syst.*, Mar. 19–21, 2008, pp. 123–127.

[9] H. Godrich, A. M. Haimovich, and R. S. Blum, "Target localisation techniques and tools for multiple-input multiple-output radar," *IET Radar, Sonar Navigat.*, vol. 3, no. 4, pp. 314–327, Aug. 2009.

[10] Q. He, R. S. Blum, H. Godrich, and A. M. Haimovich, "Target velocity estimation and antenna placement for MIMO radar with widely separated antennas," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 79–100, Feb. 2010.

[11] H. Godrich, A. M. Haimovich, and R. S. Blum, "Target localization accuracy gain in MIMO radar-based systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2783–2803, Jun. 2010.

[12] H. Godrich, A. P. Petropulu, and H. V. Poor, "Power allocation strategies for target localization in distributed multiple-radar architectures," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3226–3240, Jul. 2011.

[13] N. Garcia, A. M. Haimovich, M. Coulon, and M. Lops, "Resource allocation in MIMO radar with multiple targets for non-coherent localization," *IEEE Trans. Signal Process.*, vol. 62, no. 10, pp. 2656–2666, May 2014.

[14] N. H. Lehmann *et al.*, "Evaluation of transmit diversity in MIMO-radar direction finding," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2215–2225, May 2007.

[15] C. Wei, Q. He, and R. S. Blum, "Cramer-Rao bound for joint location and velocity estimation in multi-target non-coherent MIMO radars," in *Proc. 44th Annu. Conf. Inf. Sci. Syst.*, Mar. 17–19, 2010, pp. 1–6.

[16] Q. He and R. S. Blum, "CramerRao bound for MIMO radar target localization with phase errors," *IEEE Signal Process. Lett.*, vol. 17, no. 1, pp. 83–86, Jan. 2010.

[17] A. J. Conejo, E. Castillo, R. Minguez, and R. Garcia-Bertrand, *Decomposition Techniques in Mathematical Programming: Engineering and Science Applications*. Berlin, Germany: Springer-Verlag, 2006.

[18] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[19] A. Simonetto and G. Leus, "Distributed maximum likelihood sensor network localization," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1424–1437, Mar. 2014.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[21] H. Godrich, A. M. Haimovich, and R. S. Blum, "Cramer Rao bound on target localization estimation in MIMO radar systems," in *Proc. 42nd Annu. Conf. Inf. Sci. Syst.*, Mar. 19–21, 2008, pp. 134–139.

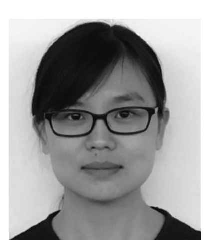
[22] H. Godrich, A. P. Petropulu, and H. V. Poor, "Sensor selection in distributed multiple-radar architectures for localization: A knapsack problem formulation," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 247–260, Jan. 2012.

[23] F. Yousefian, A. Nedić, and U. V. Shanbhag, "On stochastic gradient and subgradient methods with adaptive steplength sequences," *Automatica*, vol. 48, no. 1, pp. 56–67, Jan. 2012.

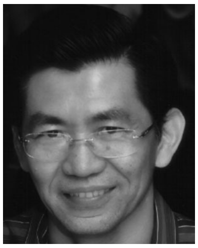
[24] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.

[25] S. Magnússon, P. C. Weeraddana, M. G. Rabbat, and C. Fischione, "On the convergence of alternating direction Lagrangian methods for nonconvex structured optimization problems," *IEEE Trans. Control Netw. Syst.*, vol. PP, no. 99, pp. 1–14, doi: 10.1109/TCNS.2015.2476198, preprint. Q2

[26] B. T. Polyak, *Introduction to Optimization*. New York, NY, USA: Optimization Software, Inc., 1987.

**Ying Ma** received the B.E. degree in electronic engineering from the Changchun Institute of Technology, Jin Lin, China, in 2011. She is currently working toward the Ph.D. degree with the Research Institute of Communication Technology, Beijing Institute of Technology, Beijing, China. Since July 2015, she has been with the Department of Electrical and Computer Engineering, University of Southampton, Southampton, U.K., where she is a visiting Ph.D. student under the supervision of Prof. L. Hanzo and Prof. S. Chen. Her research interests include distributed computation, multiple-input multiple-output systems, and cooperative communication.



**Sheng Chen** (M'90–SM'97–F'08) received the B.Eng. degree in control engineering from the East China Petroleum Institute, Dongying, China, in January 1982, the Ph.D. degree in control engineering from the City University, London, U.K., in September 1986, and the D.Sc. degree from the University of Southampton, Southampton, U.K., in 2005. He held research and academic appointments with the University of Sheffield, the University of Edinburgh, and the University of Portsmouth, U.K., from 1986 to 1999. Since 1999, he has been with the Electronics and Computer Science Department, University of Southampton, where he is a Professor of intelligent systems and signal processing. His research interests include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, evolutionary computation methods, and optimization. He has authored more than 550 research papers. He is a Fellow of the United Kingdom Royal Academy of Engineering, a Fellow of the IET, and a Distinguished Adjunct Professor with King Abdulaziz University, Jeddah, Saudi Arabia. He was an ISI Highly Cited Researcher in engineering in 2004.

**Chengwen Xing** (S'08–M'10) received the B.Eng. degree from Xidian University, Xi'an, China, in 2005, and the Ph.D. degree from the University of Hong Kong, Hong Kong, in 2010. Since September 2010, he has been with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China, where he is currently an Associate Professor. From September 2012 to December 2012, he was a Visiting Scholar at the University of Macau. His current research interests include statistical signal processing, convex optimization, multivariate statistics, combinatorial optimization, massive MIMO systems, and high-frequency band communication systems. He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *KSII Transactions on Internet and Information Systems*, *Transactions on Emerging Telecommunications Technologies*, and *China Communications*.

**Xiangyuan Bu** received the B.Eng., Master's, and Ph.D. degrees from the Beijing Institute of Technology, Beijing, China, in 1987, 1993, and 2007, respectively. He is currently a Professor in the Wireless Communications and Networks Laboratory, School of Information and Electronics, Beijing Institute of Technology. From July 1987 to September 1990, he was a Researcher in the Institute of Harbin, 674 Factory, and from April 1993 to May 2002, he was a Researcher in the No. 23 Institute of China Aerospace Science and Industry. His current research interests include wireless communication and signal processing.

**Lajos Hanzo** (F'04) received the D.Sc. degree in electronics in 1976 and the Doctorate degree in 1983. During his 38-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, U.K., where he holds the Chair in telecommunications. He has successfully supervised 100+ Ph.D. students, coauthored 20 John Wiley/IEEE Press books on mobile radio communications, totalling in excess of 10 000 pages, published 1590 research entries at IEEE Xplore, acted both as the TPC Chair and the General Chair of IEEE conferences, presented keynote lectures, and has received a number of distinctions. He is currently directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council, U.K., the European Research Council's Advanced Fellow Grant, and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. In 2009, he received the honorary doctorate "Doctor Honoris Causa" by the Technical University of Budapest. During 2008–2012, he was the Editor-in-Chief of the IEEE Press and also a Chaired Professor with Tsinghua University, Beijing, China. His research is funded by the European Research Council's Senior Research Fellow Grant. He is a Fellow of the Royal Academy of Engineering, IET, and EURASIP. For further information on research in progress and associated publications, please refer to http://www-mobile.ecs.soton.ac.uk. He has 25 000+ citations.

**QUERIES**

Q1. Author: “Scenario” is spelled as “senario.” Please check.
Q2. Author: Please provide full bibliographic details in Ref. [25].

# Decomposition Optimization Algorithms for Distributed Radar Systems

Ying Ma, Sheng Chen, *Fellow, IEEE*, Chengwen Xing, *Member, IEEE*, Xiangyuan Bu, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—Distributed radar systems are capable of enhancing the detection performance by using multiple widely spaced distributed antennas. With prior statistic information of targets, resource allocation is of critical importance for further improving the system's achievable performance. In this paper, the total transmitted power is minimized at a given mean-square target-estimation error. We derive two iterative decomposition algorithms for solving this nonconvex constrained optimization problem, namely, the optimality condition decomposition (OCD)-based and the alternating direction method of multipliers (ADMM)-based algorithms. Both the convergence performance and the computational complexity of our algorithms are analyzed theoretically, which are then confirmed by our simulation results. The OCD method imposes a much lower computational burden per iteration, while the ADMM method exhibits a higher per-iteration complexity, but as a benefit of its significantly faster convergence speed, it requires less iterations. Therefore, the ADMM imposes a lower total complexity than the OCD. The results also show that both of our schemes outperform the state-of-the-art benchmark scheme for the multiple-target case, in terms of the total power allocated, at the cost of some degradation in localization accuracy. For the single-target case, all the three algorithms achieve similar performance. Our ADMM algorithm has similar total computational complexity per iteration and convergence speed to those of the benchmark.**

***Index Terms*—Alternating direction method of multipliers, localization, multiple-input multiple-output radar, optimality condition decomposition, resource allocation.**

## I. Introduction

MULTIPLE-input multi-output (MIMO) radar systems relying on widely-separated antennas have attracted considerable attention from both industry and academia. The family of distributed MIMO radar systems is capable of significantly improving the estimation/detection performance [1]–[6] by exploiting the increased degrees of freedom resulting from the improved spatial diversity. In particular, distributed radar systems are capable of improving accuracy of target location and velocity estimation by exploiting the different Doppler estimates from multiple spatial directions [7]–[10].

Naturally, the localization performance of MIMO radar systems relying on widely-spaced distributed antennas, quantified in terms of the mean square estimation error (MSE), is determined by diverse factors, including effective signal bandwidth, the signal-to-noise ratio (SNR), the product of the numbers of transmit and receive antennas, etc [11]. Since the SNR is influenced by the path loss, the target radar cross section (RCS) and the transmitted power, the attainable localization performance can be improved by increasing either the number of participating radars or the transmitted power. However, simply increasing the amount of resources without considering the cooperation among the individual terminals is usually far from optimal.

In most traditional designs, the system's power budget is usually allocated to the transmit radars and it is fixed [6], [10], which is easy to implement and results in the simplest network structure. However, when prior estimation of the target RCS is available, according to estimation theory, uniform power allocation is far from the best strategy. In battlefields, a radar system is usually supported by power-supply trucks, but under hostile environments, their number is strictly limited. Thus, how to allocate limited resources to multiple radar stations is of great importance for maximizing the achievable performance. In other words, power allocation substantially affects the detection performance of multi-radar systems.

Recently, various studies used the Cramer-Rao lower bound (CRLB) for evaluating the performance of MIMO radar systems [11]–[16]. A power allocation scheme [12] based on CRLB was designed for multiple radar systems with a single target. The resultant nonconvex optimization problem was solved by relaxation and a domain-decomposition method. Specifically, in [12] the total transmitted power was minimized at a given estimation MSE threshold. However the algorithm of [12] was not designed for multiple-target scenarios, which are often encountered in practice. In [13] a power allocation algorithm was proposed for the multiple-target case, which is equally applicable to the single-target senario.

Against this background, in this paper, we propose two iterative decomposition methods, which are referred to as the optimality condition decomposition (OCD) [17] and the alternating direction method of multipliers (ADMM) [18] algorithms, in order to minimize the total transmitted power while satisfying a predefined estimation MSE threshold. These two algorithms can be applied to both multiple-target and single-target scenarios. The ADMM method has been widely adopted for solving convex problems. In this paper, we extend the ADMM algorithm to nonconvex problems and show that it is capable of converging.

Manuscript received March 8, 2016; revised June 17, 2016 and July 18, 2016; accepted August 9, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Fauzia Ahmad. This work was supported in part by the National Natural Science Foundation of China under Grants 61421001 and 61671058.

Y. Ma, C. Xing, and X. Bu are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: mayingbit2011@gmail.com; xingchengwen@gmail.com; bxy@bit.edu.cn).

S. Chen is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K., and also with King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: sqc@ecs.soton.ac.uk).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2016.2602801

It is worth pointing out that Simonetto and Leus [19] applied the ADMM method to solve a localization problem in a sensor network by converting the nonconvex problem to a convex one using rank-relaxation. However, the algorithm of [19] cannot be applied to our problem, because the task of [19] is that of locating sensors, which is not directly related to the signal waveform and power. Furthermore, the maximum likelihood (ML) criterion can be used for solving this sensor localization problem. However, our task is to assign the power of every MIMO radar transmitter, and at the time of writing it is an open challenge to design the ML estimator for this task [11]. The main contributions of our work are as follows.

- We propose two iterative decomposition algorithms, namely, the OCD-based and ADMM-based methods, for both multiple-target and single-target scenarios. The convergence of these two algorithms is analyzed theoretically and verified by simulations. Both these two methods are capable of converging to locally optimal solutions. The complexity analysis of the two algorithms is provided and it is shown that the OCD method imposes a much lower computational burden per iteration, while the ADMM method enjoys a significantly faster convergence speed and therefore it actually imposes a lower total complexity.
- In the multiple-target case, we demonstrate that both of our two algorithms outperform the state-of-the-art benchmark scheme of [13], in terms of the total power allocated at the expense of some degradation in localization accuracy. We show furthermore that our ADMM-based algorithm and the algorithm of [13] have similar convergence speed and total computational complexity.
- In the single-target case, we show that all the three methods attain a similar performance, since the underlying optimization problems are identical. We also prove that the closed-form solution of [12] is invalid for the systems with more than three transmit radars and we propose a beneficial suboptimal closed-form solution.

The paper is organized as follows. In Section II, the MIMO radar system model is introduced and the corresponding optimization problem is formulated. Our power allocation strategies are proposed in Section III for both the multiple-target and single-target cases, while our convergence and complexity analysis is provided in Section IV. Section V presents our simulation results for characterizing the attainable performance of the proposed algorithms which are then compared to the scheme of [13]. Finally, our conclusions are offered in Section VI.

Throughout our discussions, the following notational conventions are used. Boldface lower- and upper-case letters denote vectors and matrices, respectively. The transpose, conjugate and inverse operators are denoted by $(\cdot)^{\mathrm{T}}$, $(\cdot)^{*}$ and $(\cdot)^{-1}$, respectively, while $\mathrm{Tr}\,(\cdot)$ stands for the matrix trace operation and $\mathrm{diag}\,(x_1, x_2, \cdots, x_n)$ or $\mathrm{diag}(\mathbf{x})$ is the diagonal matrix with the specified diagonal elements. Additionally, $\mathrm{diag}\,(\mathbf{X}_1, \cdots, \mathbf{X}_K)$ and $\mathrm{diag}\,(\mathbf{x}_1, \cdots, \mathbf{x}_K)$ denotes the block diagonal matrices with the specified sub-matrices and vectors, respectively, at the corresponding block diagonal positions. The operator $\mathrm{v}_{\mathrm{diag}}(\mathbf{X})$ forms a vector using the diagonal elements of square matrix $\mathbf{X}$, while $\mathrm{E}\{\cdot\}$ denotes the expectation operator and $\otimes$ is the

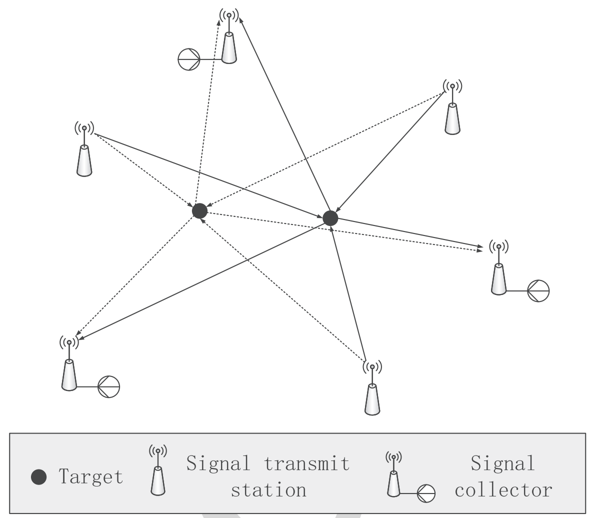



Fig. 1. Illustration of distributed radar network.

Kronecker product operator. The sub-matrix consisting of the elements of the $i_1$ to $i_2$ rows and $j_1$ to $j_2$ columns of $\mathbf{A}$ is denoted by $[\mathbf{A}]_{[i_1:i_2;j_1:j_2]}$, and the $i$th row and $j$th column element of $\mathbf{A}$ is given by $[\mathbf{A}]_{i,j}$. Similarly, $[\mathbf{a}]_{[i_1:i_2]}$ is the vector consisting of $i_1$th to $i_2$th elements of $\mathbf{a}$. The magnitude operator is given by $|\cdot|$, and $\|\cdot\|$ denotes the vector two-norm or matrix Frobenius norm. $\mathbf{I}_K$ is the identity matrix of size $K \times K$ and $\mathbf{0}$ is the zero matrix/vector of an appropriate size, while $\mathbf{1}$ denotes the vector of an appropriate size, whose elements are all equal to one. Finally, $\Re[\,]$ denotes the real part of a complex value and $\mathrm{j} = \sqrt{-1}$ represents the imaginary axis.

## II. System Model

The MIMO radar system consists of $M$ transmit radars and $N$ receive radars which cooperate to locate $K$ targets, as illustrated in Fig. 1. The $M$ transmit radars are positioned at the coordinates $(x_m^{tx}, y_m^{tx})$ for $1 \leq m \leq M$, and the $N$ receive radars are positioned at $(x_n^{rx}, y_n^{rx})$ for $1 \leq n \leq N$, while the position of target $k$ is $(x_k, y_k)$. A set of mutually orthogonal waveforms are transmitted from the transmit radars, and the corresponding baseband signals are denoted by $\left\{s_m(t)\right\}_{m=1}^{M}$ with normalized power, i.e., $\int_{\tau_m} |s_m(t)|^2 \, dt = 1$, where $\tau_m$ is the duration of the $m$th transmitted signal. Furthermore, the orthogonality of the transmitted waveforms can always be guaranteed even for different time delays, i.e., $\int_{\tau_m} s_m(t) s_{m'}^{*}(t-\tau)\, dt = 0$ for $m' \neq m$. The narrowband signals of the transmitted waveforms have the effective bandwidth $\beta_m$ specified by

$$\beta_m^2 = \frac{\int_W f^2 \left|S_m(f)\right|^2 \, df}{\int_W \left|S_m(f)\right|^2 \, df} \, (\mathrm{Hz})^2, \tag{1}$$

where $W$ is the frequency range of the signals, and $S_m(f)$ is the Fourier transform of $s_m(t)$ transmitted from the $m$th transmit radar. The transmitted powers of the different antennas, denoted by $\mathbf{p} = [p_1\, p_2 \cdots p_M]^{\mathrm{T}}$, are constrained by their corresponding minimum and maximum values specified by

$$\mathbf{p}_{\min} = \left[p_{1_{\min}}\, p_{2_{\min}} \cdots p_{M_{\min}}\right]^{\mathrm{T}}, \tag{2}$$

$$\mathbf{p}_{\max} = \left[p_{1_{\max}}\, p_{2_{\max}} \cdots p_{M_{\max}}\right]^{\mathrm{T}}. \tag{3}$$

The upper bound $p_{m_{\max}}$ is determined by the design and the lower bound $p_{m_{\min}}$ is used to guarantee that the transmit radar $m$ operates at an appropriate SNR. Let the propagation path spanning from the transmitter $m$ to the target $k$ and from the target $k$ to the receiver $n$ be defined as the channel $(m,k,n)$. Then the propagation time $\tau_{m,n}^{(k}$ of the channel $(m,k,n)$ can be calculated by $\tau_{m,n}^{(k} = (R_{m,k}^{tx} + R_{n,k}^{rx})/c$, where $c$ is the speed of light, $R_{m,k}^{tx} = \sqrt{(x_m^{tx} - x_k)^2 + (y_m^{tx} - y_k)^2}$ is the distance from transmitter $m$ to target $k$, and $R_{n,k}^{rx} = \sqrt{\left(x_n^{rx} - x_k\right)^2 + \left(y_n^{rx} - y_k\right)^2}$ is the distance from target $k$ to receiver $n$. The time delay $\tau_{m,n}^{(k}$ is used to estimate the position of targets. For far field signals, by retaining only the linear terms of its Taylor expansion, $\tau_{m,n}^{(k}$ can be approximated as a linear function of $x_k$ and $y_k$

$$\tau_{m,n}^{(k} \simeq -\frac{x_k}{c}\left(\cos\theta_m^{(k} + \cos\varphi_n^{(k}\right) - \frac{y_k}{c}\left(\sin\theta_m^{(k} + \sin\varphi_n^{(k}\right), \tag{4}$$

where $\theta_m^{(k}$ is the bearing angle of the transmitting radar $m$ to the target $k$ and $\varphi_n^{(k}$ is the bearing angle of the receiving radar $n$ to the target $k$, both measured with respect to the $x$ axis.

Let the complex-valued reflectivity coefficient $h_{m,n}^{(k}$ represent the attenuation and phase rotation of channel $(m,k,n)$. The baseband signal at receive radar $n$ can be expressed as

$$r_n(t) = \sum_{k=1}^{K}\sum_{m=1}^{M}\sqrt{p_m}\, h_{m,n}^{(k} s_m\left(t - \tau_{m,n}^{(k}\right) + \omega_n(t), \tag{5}$$

where $\omega_n(t)$ is a circularly symmetric complex Gaussian white noise, which is bandlimited to the system bandwidth $W$ and hence has a zero mean and $E\{|\omega_n(t)|^2\} = \sigma^2$. In our work, the path-loss $\kappa_{m,n}^{(k}$ is chosen as

$$\kappa_{m,n}^{(k} \propto \frac{1}{\left(R_{m,k}^{tx}\right)^2\left(R_{n,k}^{rx}\right)^2}. \tag{6}$$

Thus, given the complex target RCS $\zeta_{m,n}^{(k}$, the channel coefficient $h_{m,n}^{(k}$ is given by

$$h_{m,n}^{(k} = \zeta_{m,n}^{(k}\sqrt{\kappa_{m,n}^{(k}} = h_{m,n}^{(k,\mathrm{Re}} + \mathrm{j}h_{m,n}^{(k,\mathrm{Im}}, \tag{7}$$

where $h_{m,n}^{(k,\mathrm{Re}}$ and $h_{m,n}^{(k,\mathrm{Im}}$ are the real and imaginary parts of $h_{m,n}^{(k}$. Let us collect all the channel coefficients associated with the target $k$ in the $(2MN \times 1)$-element real-valued vector as

$$\mathbf{h}_k = \left[h_{1,1}^{(k,\mathrm{Re}} \cdots h_{1,N}^{(k,\mathrm{Re}} \cdots h_{M,N}^{(k,\mathrm{Re}}\, h_{1,1}^{(k,\mathrm{Im}} \cdots h_{1,N}^{(k,\mathrm{Im}} \cdots h_{M,N}^{(k,\mathrm{Im}}\right]^{\mathrm{T}}. \tag{8}$$

Similarly, we introduce the $(NM \times 1)$-element real vectors

$$\left|\mathbf{h}^{(k}\right|^2 = \left[\left|h_{1,1}^{(k}\right|^2 \cdots \left|h_{1,N}^{(k}\right|^2 \cdots \left|h_{M,1}^{(k}\right|^2 \cdots \left|h_{M,N}^{(k}\right|^2\right]^{\mathrm{T}}, \tag{9}$$

$$\left|\mathbf{h}^{(k}\right| = \left[\left|h_{1,1}^{(k}\right| \cdots \left|h_{1,N}^{(k}\right| \cdots \left|h_{M,1}^{(k}\right| \cdots \left|h_{M,N}^{(k}\right|\right]^{\mathrm{T}}. \tag{10}$$

Upon defining $\mathbf{h} = \left[\mathbf{h}_1^{\mathrm{T}}\, \mathbf{h}_2^{\mathrm{T}} \cdots \mathbf{h}_K^{\mathrm{T}}\right]^{\mathrm{T}}$ and the location vector of the $K$ targets as $\boldsymbol{l}_{x,y} = \left[x_1\, y_1 \cdots x_K\, y_K\right]^{\mathrm{T}}$, all the system's parameters can be stacked into a single real-valued vector

$$\mathbf{u} = \left[\boldsymbol{l}_{x,y}^{\mathrm{T}}\, \mathbf{h}^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{11}$$

Since the received signal (5) is also a function of the time delays $\tau_{m,n}^{(k}$, we also define the following system parameter vector

$$\boldsymbol{\psi} = \left[\tau_{1,1}^{(1} \cdots \tau_{1,N}^{(1} \cdots \tau_{M,N}^{(K}\, \mathbf{h}^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{12}$$

There exists a clear one-to-one relationship between $\mathbf{u}$ and $\boldsymbol{\psi}$.

Let $f(\mathbf{r}|\mathbf{u})$ be the conditional probability density function (PDF) of the observation vector $\mathbf{r} = [r_1(t), r_2(t), \cdots, r_N(t)]$ conditioned on $\mathbf{u}$. Similarly, we have the conditional PDF of $\mathbf{r}$ conditioned on $\boldsymbol{\psi}$. Then the unbiased estimate $\widehat{\mathbf{u}}$ of $\mathbf{u}$ satisfies the following inequality [20]

$$\mathrm{E}\left\{\left(\widehat{\mathbf{u}} - \mathbf{u}\right)\left(\widehat{\mathbf{u}} - \mathbf{u}\right)^{\mathrm{T}}\right\} \geq \mathbf{J}^{-1}(\mathbf{u}), \tag{13}$$

where the Fisher information matrix (FIM) $\mathbf{J}(\mathbf{u})$ is defined by

$$\mathbf{J}(\mathbf{u}) = \mathrm{E}\left\{\frac{\partial}{\partial\mathbf{u}}\log f(\mathbf{r}|\mathbf{u})\left(\frac{\partial}{\partial\mathbf{u}}\log f(\mathbf{r}|\mathbf{u})\right)^{\mathrm{T}}\right\}. \tag{14}$$

Similarly, we have the FIM of $\boldsymbol{\psi}$, denoted by $\mathbf{J}(\boldsymbol{\psi})$. The FIM $\mathbf{J}(\mathbf{u})$ can be derived from $\mathbf{J}(\boldsymbol{\psi})$ according to

$$\mathbf{J}(\mathbf{u}) = \begin{bmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{2KMN}\end{bmatrix}\mathbf{J}(\boldsymbol{\psi})\begin{bmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{2KMN}\end{bmatrix}^{\mathrm{T}}, \tag{15}$$

where the $(2K \times KMN)$-element block diagonal matrix $\mathbf{D}$ takes the following form

$$\mathbf{D} = \mathrm{diag}\left(\mathbf{D}^{(1}, \mathbf{D}^{(2}, \cdots, \mathbf{D}^{(K}\right), \tag{16}$$

with the $(2 \times MN)$-element sub-matrix $\mathbf{D}^{(k}$ given by

$$\mathbf{D}^{(k} = \begin{bmatrix}\dfrac{\partial\tau_{1,1}^{(k}}{\partial x_k} & \cdots & \dfrac{\partial\tau_{M,N}^{(k}}{\partial x_k}\\ \dfrac{\partial\tau_{1,1}^{(k}}{\partial y_k} & \cdots & \dfrac{\partial\tau_{M,N}^{(k}}{\partial y_k}\end{bmatrix} = -\frac{1}{c}\begin{bmatrix}\cos\left(\theta_1^{(k)}\right)+\cos\left(\varphi_1^{(k)}\right) & \cdots & \cos\left(\theta_M^{(k)}\right)+\cos\left(\varphi_N^{(k)}\right)\\ \sin\left(\theta_1^{(k)}\right)+\sin\left(\varphi_1^{(k)}\right) & \cdots & \sin\left(\theta_M^{(k)}\right)+\sin\left(\varphi_N^{(k)}\right)\end{bmatrix}. \tag{17}$$

The matrix $\mathbf{C}_{x,y}$ associated with the CRLB for the unbiased estimator of $\boldsymbol{l}_{x,y}$ is the $(2K \times 2K)$-element upper left block sub-matrix of $\mathbf{J}^{-1}(\mathbf{u})$, which can be derived as [11], [21]

$$\mathbf{C}_{x,y} = \left[\mathbf{J}^{-1}(\mathbf{u})\right]_{[1:2K;1:2K]} = \left(\mathbf{D}\mathbf{P}\boldsymbol{\Psi}\mathbf{D}^{\mathrm{T}}\right)^{-1}, \tag{18}$$

where $\mathbf{P} = \mathbf{I}_K \otimes \mathrm{diag}(\mathbf{p}) \otimes \mathbf{I}_N$, and $\boldsymbol{\Psi} = \mathrm{diag}\big(\boldsymbol{\Psi}^{(1}, \cdots, \boldsymbol{\Psi}^{(K)}\big)$ is the $(KMN \times KMN)$-element block diagonal matrix with the $k$th sub-matrix defined as

$$\boldsymbol{\Psi}^{(k} = 8\pi^2 \left(\mathrm{diag}\left(\beta_1^2, \cdots, \beta_M^2\right) \otimes \mathbf{I}_N\right) \mathrm{diag}\left(\left|\mathbf{h}^{(k}\right|^2\right). \quad (19)$$

Let us denote the variances of the estimates of $x_k$ and $y_k$ by $\sigma_{x_k}^2$ and $\sigma_{y_k}^2$, respectively. Then we have

$$\sum_{k=1}^{K} \left(\sigma_{x_k}^2 + \sigma_{y_k}^2\right) \geq \mathrm{Tr}\left(\mathbf{C}_{x,y}\right), \quad (20)$$

where $\mathrm{Tr}\left(\mathbf{C}_{x,y}\right)$ is a lower bound on the sum of the MSEs of the localization estimator $\widehat{\boldsymbol{l}}_{x,y}$. By defining $\mathbf{X} = \mathrm{diag}(\mathbf{p}) \otimes \mathbf{I}_N$ and noting $\mathbf{D}$ of (16), we obtain the expression of the lower bound for the $k$th target location estimate as [12], [22]

$$\begin{aligned} &\sum_{i=1}^{2} \left[\mathbf{C}_{x,y}\right]_{i+2(k-1),i+2(k-1)} \\ &= \sum_{i=1}^{2} \left[\left(\mathbf{D}\mathbf{P}\boldsymbol{\Psi}\mathbf{D}^{\mathrm{T}}\right)^{-1}\right]_{i+2(k-1),i+2(k-1)} \\ &= \mathrm{Tr}\left(\begin{bmatrix} \left(\mathbf{a}_{1,1}^{(k}\right)^{\mathrm{T}}\mathbf{p} & \left(\mathbf{a}_{1,2}^{(k}\right)^{\mathrm{T}}\mathbf{p} \\ \left(\mathbf{a}_{2,1}^{(k}\right)^{\mathrm{T}}\mathbf{p} & \left(\mathbf{a}_{2,2}^{(k}\right)^{\mathrm{T}}\mathbf{p} \end{bmatrix}^{-1}\right) = \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}}, \end{aligned} \quad (21)$$

where the second equation is obtained by first dividing the $(MN \times 2)$ matrix $\left(\mathbf{D}^{(k)}\right)^{\mathrm{T}}$ into the two column vectors, $\left(\mathbf{D}^{(k)}\right)^{\mathrm{T}} = \left[\mathbf{d}_1^{(k}\ \mathbf{d}_2^{(k}\right]$, and generating the $(N \times 1)$ vectors

$$\mathbf{d}_{i,m}^{(k} = \left[\mathbf{d}_i^{(k}\right]_{[(m-1)N+1:mN]}, \ i = 1, 2, \ 1 \leq m \leq M. \quad (22)$$

Then $\mathbf{a}_{i,j}^{(k}$ for $1 \leq i, j \leq 2$ are given by

$$\begin{aligned} \mathbf{a}_{i,j}^{(k} = \mathrm{v}_{\mathrm{diag}} &\left(\mathrm{diag}\left(\left(\mathbf{d}_{i,1}^{(k}\right)^{\mathrm{T}}, \cdots, \left(\mathbf{d}_{i,M}^{(k}\right)^{\mathrm{T}}\right) \boldsymbol{\Psi}^{(k}\right. \\ &\left.\times \mathrm{diag}\left(\mathbf{d}_{j,1}^{(k}, \cdots, \mathbf{d}_{j,M}^{(k}\right)\right), \end{aligned} \quad (23)$$

while $\mathbf{b}_k = \mathbf{a}_{1,1}^{(k} + \mathbf{a}_{2,2}^{(k}$ and $\mathbf{A}_k = \mathbf{a}_{1,1}^{(k}\left(\mathbf{a}_{2,2}^{(k}\right)^{\mathrm{T}} - \mathbf{a}_{1,2}^{(k}\left(\mathbf{a}_{2,1}^{(k}\right)^{\mathrm{T}}$.

Our task is to design a beneficial power allocation strategy capable of achieving a localization accuracy threshold $\eta$. We can use the weighting $v_k$ to indicate the localization accuracy requirement for the $k$th target. The larger $v_k$ is, the higher accuracy is required for the $k$th target. For a predetermined lower bound of total MSE of all the targets, the transmit power of the different transmit radars can then be appropriately allocated for minimizing the total transmit power. This can be formulated as the following optimization problem $\mathbb{P}1$

$$\mathbb{P}1: \begin{array}{ll} \min\limits_{\mathbf{p}} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & \sum\limits_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}} \leq \eta, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \ 1 \leq m \leq M. \end{array} \quad (24)$$

Because generally speaking $\mathbf{A}_k$ is not a positive definite matrix, the optimization $\mathbb{P}1$ is a nonconvex problem.

In [13], a similar optimization problem is formulated as

$$\begin{array}{ll} \min\limits_{\mathbf{p}} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}} \leq \bar{\eta}, \ 1 \leq k \leq K, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \ 1 \leq m \leq M, \end{array} \quad (25)$$

given an equivalent localization accuracy threshold $\bar{\eta}$. In [13], a Taylor series based technique is applied to approximate the inequality constraints in (25) in order to relax the nonconvex optimization problem for the sake of obtaining a solution. Intuitively, the cost function associated with an optimal solution of our problem $\mathbb{P}1$ of (24) is generally smaller than that associated with an optimal solution of (25), i.e., we can achieve a lower power consumption. This is achieved at the potential cost of a slightly reduced localization accuracy.

## III. Power Resource Allocation

### A. Multi-Target Case

In order to solve the nonconvex problem $\mathbb{P}1$ of (24), we have to change it into a simpler form. Specifically, we have to change the inequality constraint into an equality one, i.e.,

$$\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}} \leq \eta \Rightarrow \sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}} = \eta. \quad (26)$$

*Lemma 1:* An increase of the transmit power $\mathbf{p}$ results in a reduction of the MSE, namely,

$$\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}(\mathbf{p} + \Delta\mathbf{p})}{(\mathbf{p} + \Delta\mathbf{p})^{\mathrm{T}}\mathbf{A}_k(\mathbf{p} + \Delta\mathbf{p})} \leq \sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}}. \quad (27)$$

The proof of Lemma 1 is similar to that of single-target case given in [12]. Thus, to achieve a reduced power consumption, we can always set the MSE to its maximum tolerance. The change of constraint as given in (26) leads to the problem $\mathbb{P}2$,

$$\mathbb{P}2: \begin{array}{ll} \min\limits_{\mathbf{p}} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & \sum\limits_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}} = \eta, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \ 1 \leq m \leq M. \end{array} \quad (28)$$

*Theorem 1:* The solutions of $\mathbb{P}1$ and $\mathbb{P}2$ are identical.

The proof of Theorem 1 is straightforward. By introducing the auxiliary variables

$$w_k = \frac{1}{\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}}, \ 1 \leq k \leq K, \quad (29)$$

and their corresponding lower and upper bounds

$$w_{k_{\min}} = \frac{1}{\eta\mathbf{p}_{\max}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_{\max}}, \ w_{k_{\max}} = \frac{1}{\eta\mathbf{p}_{\min}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_{\min}}, \ 1 \leq k \leq K, \quad (30)$$

$\mathbb{P}2$ is reformulated as the following optimization problem $\mathbb{P}3$:

$$\mathbb{P}3: \quad \begin{array}{ll} \min\limits_{\mathbf{p},\mathbf{w}} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & \sum\limits_{k=1}^{K} v_k w_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} = 1, \\ & w_k \eta \mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} = 1,\ 1 \le k \le K, \\ & p_{m_{\min}} \le p_m \le p_{m_{\max}},\ 1 \le m \le M, \\ & w_{k_{\min}} \le w_k \le w_{k_{\max}},\ 1 \le k \le K. \end{array} \tag{31}$$

The following corollary is obvious.

*Corollary 1:* If $\mathbf{p}^\star$ associated with $w_k^\star = \frac{1}{\eta\left(\mathbf{p}^\star\right)^{\mathrm{T}}\mathbf{A}_k\mathbf{p}^\star}$ for $1 \le k \le K$ is an optimal solution of the problem $\mathbb{P}3$ (31), $\mathbf{p}^\star$ is an optimal solution for the problem $\mathbb{P}1$ of (24). Conversely, if $\mathbf{p}^\star$ is an optimal solution of the problem $\mathbb{P}1$, together with $w_k^\star = \frac{1}{\eta\left(\mathbf{p}^\star\right)^{\mathrm{T}}\mathbf{A}_k\mathbf{p}^\star}$ for $1 \le k \le K$ it is an optimal solution of the problem $\mathbb{P}3$.

*1) OCD-based method:* The Lagrangian associated with the optimization problem $\mathbb{P}3$ is

$$\begin{aligned} L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) = \mathbf{1}^{\mathrm{T}}\mathbf{p} + \lambda\left(\sum_{k=1}^{K} v_k w_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1\right) \\ + \sum_{k=1}^{K}\mu_k\left(w_k\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} - 1\right), \end{aligned} \tag{32}$$

with $\mathbf{w} = \left[w_1\ w_2 \cdots w_K\right]^{\mathrm{T}}$ and $\boldsymbol{\mu} = \left[\mu_1\ \mu_2 \cdots \mu_K\right]^{\mathrm{T}}$, where $\lambda$ and $\mu_k$ for $1 \le k \le K$ are Lagrangian multipliers. We optimize the Lagrangian (32) with respect to $\mathbf{p}$, $\lambda$, $w_k$ and $\mu_k$. Using the steepest descent method, the search directions are related to the Karush-Kuhn-Tucker (KKT) conditions by

$$\begin{aligned} \Delta\mathbf{p} = \nabla_{\mathbf{p}}L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) = \mathbf{1} + \lambda\left(\sum_{k=1}^{K} w_k v_k \mathbf{b}_k\right) \\ + \sum_{k=1}^{K}\mu_k w_k \eta\left(\mathbf{A}_k + \mathbf{A}_k^{\mathrm{T}}\right)\mathbf{p}, \end{aligned} \tag{33}$$

$$\Delta\lambda = \nabla_{-\lambda}L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) = -\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} + 1, \tag{34}$$

$$\begin{aligned} \Delta w_k &= \nabla_{w_k}L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) \\ &= \lambda v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} + \mu_k\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p},\ 1 \le k \le K, \end{aligned} \tag{35}$$

$$\begin{aligned} \Delta\mu_k &= \nabla_{-\mu_k}L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu}) \\ &= -\left(\eta w_k\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} + 1\right),\ 1 \le k \le K, \end{aligned} \tag{36}$$

where we have $\Delta\mathbf{p} = \left[\Delta p_1\ \Delta p_2 \cdots \Delta p_M\right]^{\mathrm{T}}$. The primal and dual variables are updated iteratively

$$p_m^{(n+1)} = \left[p_m^{(n)} - \kappa_1\Delta p_m^{(n)}\right]_{p_{m\,\min}}^{p_{m\,\max}},\ 1 \le m \le M, \tag{37}$$

$$\lambda^{(n+1)} = \lambda^{(n)} - \kappa_2\Delta\lambda^{(n)}, \tag{38}$$

$$w_k^{(n+1)} = w_k^{(n)} - \kappa_3\Delta w_k^{(n)},\ 1 \le k \le K, \tag{39}$$

$$\mu_k^{(n+1)} = \mu_k^{(n)} - \kappa_4\Delta\mu_k^{(n)},\ 1 \le k \le K, \tag{40}$$

where the superscript $^{(n)}$ denotes the iteration index and

$$[a]_b^c = \min\left\{\max\left\{a,b\right\},c\right\}, \tag{41}$$

while $\kappa_i$ for $1 \le i \le 4$ represent the step sizes for the primal variables $\mathbf{p}$, the dual variable $\lambda$, the primal variables $\mathbf{w}$ and the dual variables $\boldsymbol{\mu}$, respectively. According to [23], an exponentially decreasing step size is highly desired. Furthermore, since $\mathbf{p}$, $\lambda$, $\mathbf{w}$ and $\boldsymbol{\mu}$ have very different properties and their impacts on the Lagrangian are 'unequal', using different step sizes for them makes sense. By combining these two considerations, we set the four step sizes for $\mathbf{p}$, $\lambda$, $\mathbf{w}$ and $\boldsymbol{\mu}$ according to

$$\kappa_i = c_i e^{-\alpha n} \text{ with } 0 \le \alpha \ll 1,\ \text{for } 1 \le i \le 4, \tag{42}$$

where $c_i > 0$ for $1 \le i \le 4$ are different constants.

The choice of the initial values for the primal variables $p_m$, $1 \le m \le M$, influences the convergence performance. Ideally, they should be chosen to be close to their own specific optimal values so as to enhance the convergence speed. For practical reason, the initialization should be easy and simple to realize too. Hence we opt for the initial powers of

$$\mathbf{p}^{(0)} = \mathbf{p}_{equ} = \frac{1}{\eta}\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{A}_k\mathbf{1}}\mathbf{1}, \tag{43}$$

which is obtained by setting all the elements of $\mathbf{p}$ to be equal. Then, $w_k$ is initialized according to

$$w_k^{(0)} = \frac{1}{\eta\mathbf{p}_{equ}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_{equ}}, 1 \le k \le K. \tag{44}$$

The iterative procedure of (37) to (40) is repeated until $\left\|\mathbf{p}^{(n+1)} - \mathbf{p}^{(n)}\right\|$ becomes smaller than a preset small positive number or the maximum number of iterations is reached.

*Remark 1*: It is difficult to find a closed-form solution from the set of KKT conditions, because $\mathbf{A}_k$ for $1 \le k \le K$ are generally non-invertible. Hence our algorithm finds a locally optimal point in an iterative manner. It is also worth noting that the standard OCD [17] is typically based on a Newton-type algorithm, but our proposed OCD method is a steepest descent algorithm. The reason is that the Hessian matrix for the Lagrangian $L(\mathbf{p},\mathbf{w},\lambda,\boldsymbol{\mu})$ of (32) is not invertible.

*2) ADMM-based method:* ADMM was originally proposed for solving convex problems in a parallel manner [18]. Let us now discuss how to apply the ADMM method for solving the nonconvex problem $\mathbb{P}3$. By introducing an auxiliary vector $\mathbf{z} = \mathbf{p}$, (29) can be rewritten as

$$\mathbf{p} = \mathbf{z} \ \text{ and } \ \eta w_k \mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} = 1,\ 1 \le k \le K. \tag{45}$$

Therefore, $\mathbb{P}3$ can be reformulated into the problem $\mathbb{P}4$:

$$\mathbb{P}4: \quad \begin{array}{ll} \min\limits_{\mathbf{p},\mathbf{w},\mathbf{z}} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & \sum\limits_{k=1}^{K} v_k w_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} = 1, \\ & \mathbf{p} = \mathbf{z}, \\ & w_k \eta \mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p} = 1,\ 1 \le k \le K, \\ & p_{m_{\min}} \le p_m \le p_{m_{\max}},\ 1 \le m \le M, \\ & w_{k_{\min}} \le w_k \le w_{k_{\max}},\ 1 \le k \le K. \end{array} \tag{46}$$

This problem is convex with respect to $\mathbf{p}$, $\mathbf{z}$ and $w_k$, respectively. An augmented Lagrangian is constructed as follows

$$\begin{aligned} L(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_0, d_1, \mathbf{d}_2) &= \mathbf{1}^{\mathrm{T}}\mathbf{p} + \frac{\rho_0}{2}\|\mathbf{p}-\mathbf{z}\|^2 + \mathbf{d}_0^{\mathrm{T}}(\mathbf{p}-\mathbf{z}) \\ &+ \sum_{k=1}^{K}\frac{\rho_{2,k}}{2}\left|w_k\mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta - 1\right|^2 + \sum_{k=1}^{K} d_{2,k}\left(w_k\mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta - 1\right) \\ &+ \frac{\rho_1}{2}\left|\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1\right|^2 + d_1\left(\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1\right) \end{aligned} \tag{47}$$

where $\mathbf{d}_0 = \left[d_{0,1}\cdots d_{0,M}\right]^{\mathrm{T}}$, $d_1$ and $\mathbf{d}_2 = \left[d_{2,1}\cdots d_{2,K}\right]^{\mathrm{T}}$ are the dual variables corresponding to the constraints $\mathbf{p} = \mathbf{z}$, $\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} = 1$ and $w_k\mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta = 1$ for $1 \le k \le K$, respectively, while $\rho_0$, $\rho_1$ and $\boldsymbol{\rho}_2 = \left[\rho_{2,1}\cdots\rho_{2,K}\right]^{\mathrm{T}}$ are the penalty parameters. Note that the augmented Lagrangian (47) is quadratic. For convenience, we scale the dual variables as $\mathbf{e} = \frac{1}{\rho_0}\mathbf{d}_0$, $\mu = \frac{1}{\rho_1}d_1$ and $\boldsymbol{\gamma} = \left[\gamma_1\cdots\gamma_K\right]^{\mathrm{T}}$ with $\gamma_k = \frac{1}{\rho_{2,k}}d_{2,k}$ for $1 \le k \le K$. Then, from (47) we obtain the following augmented Lagrangian

$$\begin{aligned} L(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{e}, \mu, \boldsymbol{\gamma}) &= \mathbf{1}^{\mathrm{T}}\mathbf{p} + \frac{\rho_0}{2}\|\mathbf{p}-\mathbf{z}+\mathbf{e}\|^2 - \frac{\rho_0}{2}\|\mathbf{e}\|^2 \\ &+ \sum_{k=1}^{K}\frac{\rho_{2,k}}{2}\left|w_k\mathbf{z}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta - 1 + \gamma_k\right|^2 - \sum_{k=1}^{K}\frac{\rho_{2,k}}{2}\left|\gamma_k\right|^2 \\ &+ \frac{\rho_1}{2}\left|\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1 + \mu\right|^2 - \frac{\rho_1}{2}|\mu|^2. \end{aligned} \tag{48}$$

We can find the saddle point of the augmented Lagrangian (48) by minimizing the Lagrangian over the primal variables $\mathbf{p}$, $\mathbf{w}$ and $\mathbf{z}$, as well as maximizing it over the dual variables $\mathbf{e}$, $\mu$ and $\boldsymbol{\gamma}$, in an alternative way. In particular, we update the primal variables $\mathbf{p}$, $\mathbf{w}$ and $\mathbf{z}$ separately one by one. Furthermore, after the update of the dual variables $\mathbf{e}$, $\mu$ and $\boldsymbol{\gamma}$, we adjust the penalty parameters $\rho_0$, $\rho_1$ and $\boldsymbol{\rho}_2$. We now summarize our ADMM-based procedure.

*Initialization:* Let us also opt for the equal power initialization $\mathbf{p}^{(0)} = \mathbf{p}_{equ}$ of (43). The other primal variables are initialized as $w_k^{(0)} = \frac{1}{\eta\mathbf{p}_{equ}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_{equ}}$ for $1 \le k \le K$ of (44), and

$$\mathbf{z}^{(0)} = \mathbf{p}_{equ}. \tag{49}$$

The initial penalty parameters, $\rho_0^{(0)}$, $\rho_1^{(0)}$ and $\rho_{2,k}^{(0)}$ for $1 \le k \le K$, are typically set to a large positive value, say, 500. Next, the dual variables are initialized as follows. Choose $\mu^{(0)} = 1$ and $\gamma_k^{(0)} = 1$ for $1 \le k \le K$, while every element of $\mathbf{e}^{(0)}$ is set to 1 too. Then we set the iteration index $n = 0$.

*Iterative Procedure:* At the $(n+1)$th iteration, perform:

- *Step 1: Update the primal variables* $\mathbf{p}$. Upon isolating all the terms involving $\mathbf{p}$ in the Lagrangian (48), we have

$$\begin{aligned} \min_{\mathbf{p}} \mathbf{1}^{\mathrm{T}}\mathbf{p} &+ \frac{\rho_0^{(n)}}{2}\left\|\mathbf{p}-\mathbf{z}^{(n)}+\mathbf{e}^{(n)}\right\|^2 \\ &+ \frac{\rho_1^{(n)}}{2}\left|\sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p} - 1 + \mu^{(n)}\right|^2 \\ &+ \sum_{k=1}^{K}\frac{\rho_{2,k}^{(n)}}{2}\left|w_k^{(n)}\left(\mathbf{z}^{(n)}\right)^{\mathrm{T}}\mathbf{A}_k\mathbf{p}\eta - 1 + \gamma_k^{(n)}\right|^2, \\ \text{s.t. } & p_{m_{\min}} \le p_m \le p_{m_{\max}},\ 1 \le m \le M, \end{aligned} \tag{50}$$

which is a constrained convex optimization. Setting the derivative of the objective function to zero yields the $(n+1)$th estimate of $\mathbf{p}$ as follows. First compute

$$\bar{\mathbf{p}}^{(n+1)} = \left[\bar{p}_1^{(n+1)}\cdots\bar{p}_M^{(n+1)}\right]^{\mathrm{T}} = \left(\mathbf{P}_1^{(n+1)}\right)^{-1}\mathbf{p}_2^{(n+1)}, \tag{51}$$

$$\begin{aligned} \mathbf{P}_1^{(n+1)} &= \rho_0^{(n)}\mathbf{I}_M + \rho_1^{(n)}\left(\sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k\right) \\ &\quad\times\left(\sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k^{\mathrm{T}}\right) + \sum_{k=1}^{K}\rho_{2,k}^{(n)} \\ &\quad\times\left(w_k^{(n)}(\mathbf{A}_k)^{\mathrm{T}}\mathbf{z}^{(n)}\eta\right)\left(w_k^{(n)}\left(\mathbf{z}^{(n)}\right)^{\mathrm{T}}\mathbf{A}_k\eta\right)^{\mathrm{T}}, \end{aligned} \tag{52}$$

$$\begin{aligned} \mathbf{p}_2^{(n+1)} &= -\mathbf{1} + \rho_0^{(n)}\left(\mathbf{z}^{(n)}+\mathbf{e}^{(n)}\right) \\ &\quad+ \rho_1^{(n)}\left(\sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k\right)\left(1-\mu^{(n)}\right) \\ &\quad+ \rho_{2,k}^{(n)}\left(w_k^{(n)}(\mathbf{A}_k)^{\mathrm{T}}\mathbf{z}^{(n)}\eta\right)\left(1-\gamma_k^{(n)}\right). \end{aligned} \tag{53}$$

The final estimate is then given by

$$p_m^{(n+1)} = \left[\bar{p}_m^{(n+1)}\right]_{p_{m_{\min}}}^{p_{m_{\max}}},\ 1 \le m \le M. \tag{54}$$

- *Step 2: Update the primal variables* $\mathbf{w}$. The optimization involving $\mathbf{w}$ is also a constrained convex problem

$$\begin{aligned} \min_{\mathbf{w}} &\ \frac{\rho_1^{(n)}}{2}\left|\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}}\mathbf{p}^{(n+1)} - 1 + \mu^{(n)}\right|^2 \\ &+ \sum_{k=1}^{K}\frac{\rho_{2,k}^{(n)}}{2}\left|w_k\left(\mathbf{z}^{(n)}\right)^{\mathrm{T}}\mathbf{A}_k\mathbf{p}^{(n+1)}\eta - 1 + \gamma_k^{(n)}\right|^2, \\ \text{s.t. } & w_{k_{\min}} \le w_k \le w_{k_{\max}},\ 1 \le k \le K. \end{aligned} \tag{55}$$

The solution is given by

$$w_k^{(n+1)} = \left[\frac{w_{k,1}^{(n+1)}}{w_{k,2}^{(n+1)}}\right]_{w_{k_{\min}}}^{w_{k_{\max}}},\ 1 \le k \le K, \tag{56}$$

where

$$w_{k,1}^{(n+1)} = \rho_1^{(n)} v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} \left( 1 - \mu^{(n)} - \sum_{k' \neq k} v_{k'} \mathbf{b}_{k'}^{\mathrm{T}} \mathbf{p}^{(n+1)} \right) + \rho_{2,k}^{(n)} \left( \left( \mathbf{z}^{(n)} \right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta \right) \left( 1 - \gamma_k^{(n)} \right), \tag{57}$$

$$w_{k,2}^{(n+1)} = \rho_1^{(n)} \left( v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} \right)^2 + \rho_{2,k}^{(n)} \left( \left( \mathbf{z}^{(n)} \right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta \right)^2. \tag{58}$$

- *Step 3: Update the primal variables* $\mathbf{z}$. Isolating all the terms involving $\mathbf{z}$, the optimization is an unconstrained convex problem

$$\min_{\mathbf{z}} \frac{\rho_0^{(n)}}{2} \left\| \mathbf{p}^{(n+1)} - \mathbf{z} + \mathbf{e}^{(n)} \right\|^2 + \sum_{k=1}^{K} \frac{\rho_{2,k}^{(n)}}{2} \left| w_k^{(n+1)} \mathbf{z}^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1 + \gamma_k^{(n)} \right|^2. \tag{59}$$

Solving (59) yields the $(n+1)$th estimate of $\mathbf{z}$ as

$$\mathbf{z}^{(n+1)} = \left( \mathbf{Z}_1^{(n+1)} \right)^{-1} \mathbf{z}_2^{(n+1)}, \tag{60}$$

where

$$\mathbf{Z}_1^{(n+1)} = \rho_0^{(n)} \mathbf{I}_M + \sum_{k=1}^{K} \rho_{2,k}^{(n)} \left( w_k^{(n+1)} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta \right) \times \left( w_k^{(n+1)} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta \right)^{\mathrm{T}}, \tag{61}$$

$$\mathbf{z}_2^{(n+1)} = \rho_0^{(n)} \left( \mathbf{p}^{(n+1)} + \mathbf{e}^{(n)} \right) + \sum_{k=1}^{K} \rho_{2,k}^{(n)} \left( w_k^{(n+1)} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta \right) \left( 1 - \gamma_k^{(n)} \right). \tag{62}$$

- *Step 4: Update the dual variables* $\mathbf{e}$, $\mu$ *and* $\boldsymbol{\gamma}$. Maximizing the Lagrangian (48) with respect to the dual variables yields

$$\mathbf{e}^{(n+1)} = \mathbf{e}^{(n)} + \mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)}, \tag{63}$$

$$\mu^{(n+1)} = \mu^{(n)} + \sum_{k=1}^{K} w_k^{(n+1)} v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} - 1, \tag{64}$$

$$\gamma_k^{(n+1)} = \gamma_k^{(n)} + w_k^{(n+1)} \left( \mathbf{z}^{(n+1)} \right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1, \quad 1 \leq k \leq K. \tag{65}$$

- *Step 5: Update the penalty parameters* $\rho_0$, $\rho_1$ *and* $\boldsymbol{\rho}_2$. The penalty parameters are updated at the end of each iteration for the first a few iterations to speed up the convergence. At the $(n+1)$th iteration, associated with the three penalty parameters of $\rho_0^{(n)}$, $\rho_1^{(n)}$ and $\boldsymbol{\rho}_2^{(n)}$, we have three primal residuals

$$r_0^{(n+1)} = \left\| \mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)} \right\|, \tag{66}$$

$$r_1^{(n+1)} = \left| \sum_{k=1}^{K} w_k^{(n+1)} v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} - 1 \right|, \tag{67}$$

$$r_{2,k}^{(n+1)} = \left| w_k \left( \mathbf{z}^{(n+1)} \right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1 \right|, \quad 1 \leq k \leq K, \tag{68}$$

as well as three respective dual residuals

$$s_0^{(n+1)} = \left\| \rho_0^{(n)} \left( \mathbf{z}^{(n+1)} - \mathbf{z}^{(n)} \right) \right\|, \tag{69}$$

$$s_1^{(n+1)} = \left\| \mathbf{s}_{1a}^{(n+1)} \right\|, \tag{70}$$

$$s_{2,k}^{(n+1)} = \sqrt{ \left( s_{2a,k}^{(n+1)} \right)^2 + \left\| \mathbf{s}_{2b,k}^{(n+1)} \right\| }, \; 1 \leq k \leq K, \tag{71}$$

where

$$\mathbf{s}_{1a}^{(n+1)} = \mu^{(n+1)} \rho_1^{(n)} \left( \sum_{k=1}^{K} \left( w_k^{(n)} - w_k^{(n+1)} \right) v_k \mathbf{b}_k \right) + \rho_1^{(n)} \left( \sum_{k=1}^{K} w_k^{(n)} v_k \mathbf{b}_k \right) \times \left( \sum_{k=1}^{K} \left( w_k^{(n)} - w_k^{(n+1)} \right) v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} \right), \tag{72}$$

$$s_{2a,k}^{(n+1)} = \rho_{2,k}^{(n)} \left( \mathbf{z}^{(n)} \right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta \times \left( w_k^{(n+1)} \left( \mathbf{z}^{(n)} - \mathbf{z}^{(n+1)} \right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1 \right) + \gamma_k^{(n+1)} \rho_{2,k}^{(n)} \left( \left( \mathbf{z}^{(n)} - \mathbf{z}^{(n+1)} \right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta \right), \tag{73}$$

$$\mathbf{s}_{2b,k}^{(n+1)} = \rho_{2,k}^{(n)} w_k^{(n)} \eta \mathbf{A}_k^{\mathrm{T}} \mathbf{z}^{(n)} \times \left( \left( w_k^{(n)} \left( \mathbf{z}^{(n)} \right)^{\mathrm{T}} - w_k^{(n+1)} \left( \mathbf{z}^{(n+1)} \right)^{\mathrm{T}} \right) \times \mathbf{A}_k \mathbf{p}^{(n+1)} \eta \right) + \gamma_k^{(n+1)} \rho_{2,k}^{(n)} \eta \mathbf{A}_k^{\mathrm{T}} \times \left( w_k^{(n)} \mathbf{z}^{(n)} - w_k^{(n+1)} \mathbf{z}^{(n+1)} \right). \tag{74}$$

The exact definitions of the dual residuals can be found in Appendix A.

The penalty parameter $\rho_0$ is updated as follows

$$\rho_0^{(n+1)} = \begin{cases} \tau \rho_0^{(n)}, & \text{if } r_0^{(n+1)} \geq \varepsilon s_0^{(n+1)}, \\ \frac{1}{\tau} \rho_0^{(n)}, & \text{if } s_0^{(n+1)} \geq \varepsilon r_0^{(n+1)}, \\ \rho_0^{(n)}, & \text{otherwise}, \end{cases} \tag{75}$$

where the scalars $\tau > 1$ and $\varepsilon > 1$ with typical values of $\tau = 2$ and $\varepsilon = 10$. The idea behind this penalty parameter update is to balance the primal and dual residual norms $r_0^{(n+1)}$ and $s_0^{(n+1)}$, i.e., to keep $\frac{r_0^{(n+1)}}{s_0^{(n+1)}} \leq \varepsilon$ and $\frac{s_0^{(n+1)}}{r_0^{(n+1)}} \leq \varepsilon$, as they both converge to zero [18], [25]. The related dual variables are rescaled to remove the impact of changing $\rho_0$ according to

$$\mathbf{e}^{(n+1)} = \frac{\rho_0^{(n)}}{\rho_0^{(n+1)}}\mathbf{e}^{(n)}. \tag{76}$$

Similarly, $\rho_1$ is updated according to

$$\rho_1^{(n+1)} = \begin{cases} \tau\rho_1^{(n)}, & \text{if } r_1^{(n+1)} \geq \varepsilon s_1^{(n+1)}, \\ \frac{1}{\tau}\rho_1^{(n)}, & \text{if } s_1^{(n+1)} \geq \varepsilon r_1^{(n+1)}, \\ \rho_1^{(n)}, & \text{otherwise.} \end{cases} \tag{77}$$

The related dual variable is then scaled according to

$$\mu^{(n+1)} = \frac{\rho_1^{(n)}}{\rho_1^{(n+1)}}\mu^{(n)}. \tag{78}$$

Likewise, $\rho_{2,k}$ for $1 \leq k \leq K$ are updated according to

$$\rho_{2,k}^{(n+1)} = \begin{cases} \tau\rho_{2,k}^{(n)}, & \text{if } r_{2,k}^{(n+1)} \geq \varepsilon s_{2,k}^{(n+1)}, \\ \frac{1}{\tau}\rho_{2,k}^{(n)}, & \text{if } s_{2,k}^{(n+1)} \geq \varepsilon r_{2,k}^{(n+1)}, \\ \rho_{2,k}^{(n)}, & \text{otherwise,} \end{cases} \tag{79}$$

and the corresponding dual variables are rescaled as

$$\gamma_k^{(n+1)} = \frac{\rho_{2,k}^{(n)}}{\rho_{2,k}^{(n+1)}}\gamma_k^{(n)}, \; 1 \leq k \leq K. \tag{80}$$

- *Termination of the iterative procedure*. The iterative procedure is terminated when $\left\|\mathbf{p}^{(n+1)} - \mathbf{p}^{(n)}\right\|$ becomes smaller than a predefined small positive value or the preset maximum number of iterations is reached. Otherwise, set $n = n + 1$ and go to *Step 1*.

*Remark 2*: The ADMM combines the advantages of the dual ascent and the augmented Lagrangian method. The dual ascent approach deals with the complicated constraints, while the augmented Lagrangian method is capable of enhancing the convergence rate and the robustness of the algorithm.

*Remark 3*: We deal with the optimization problem (24), and in every iteration of our OCD and ADMM methods, we have a closed-form update value. By contrast, Garcia *et al.* [13] deal with the optimization problem (25), and in every iteration, an inner iterative loop is required for computing an updated value by the algorithm of [13].

### B. Single-Target Case

The target index $k$ can be dropped and then the optimization is simplified to the problem $\mathbb{P}5$

$$\mathbb{P}5: \begin{array}{ll} \min\limits_{\mathbf{p}} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & \frac{\mathbf{b}^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p}} \leq \eta, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \tag{81}$$

In the single-target case, the optimization (25) is identical to the problem $\mathbb{P}5$. Similar to the multi-target case, the problem $\mathbb{P}5$ is equivalent to the optimization problem $\mathbb{P}6$:

$$\mathbb{P}6: \begin{array}{ll} \min\limits_{\mathbf{p},w} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & w\mathbf{b}^{\mathrm{T}}\mathbf{p} - 1 = 0, \\ & w\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - 1 = 0, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \tag{82}$$

This problem is nonconvex due to its equality constraint.

*1) OCD-based method:* The Lagrangian of (82) is

$$L(\mathbf{p}, w, \lambda, \mu) = \mathbf{1}^{\mathrm{T}}\mathbf{p} + \lambda\left(w\mathbf{b}^{\mathrm{T}}\mathbf{p} - 1\right) + \mu\left(w\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - 1\right), \tag{83}$$

where $\lambda$ and $\mu$ are the dual variables. The gradients of this Lagrangian are given by

$$\Delta\mathbf{p} = \nabla_{\mathbf{p}}L(\mathbf{p}, w, \lambda, \mu) = \mathbf{1} + \lambda\left(w\mathbf{b}\right) + \mu w\eta\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)\mathbf{p}, \tag{84}$$

$$\Delta\lambda = \nabla_{-\lambda}L(\mathbf{p}, w, \lambda, \mu) = -w\mathbf{b}^{\mathrm{T}}\mathbf{p} + 1, \tag{85}$$

$$\Delta w = \nabla_{w}L(\mathbf{p}, w, \lambda, \mu) = \lambda\mathbf{b}^{\mathrm{T}}\mathbf{p} + \mu\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p}, \tag{86}$$

$$\Delta\mu = \nabla_{-\mu}L(\mathbf{p}, w, \lambda, \mu) = -\eta w\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - 1, \tag{87}$$

Given $\lambda^{(0)}$, $\mu^{(0)}$ and

$$\mathbf{p}^{(0)} = \mathbf{p}_{equ} = \frac{1}{\eta}\frac{\mathbf{b}^{\mathrm{T}}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{A}\mathbf{1}}\mathbf{1}, \tag{88}$$

$\mathbf{p}$, $\lambda$, $w$, $\mu$ are updated in the following iterative procedure

$$p_m^{(n+1)} = \left[p_m^{(n)} - \kappa_1\Delta p_m^{(n)}\right]_{p_{m_{\min}}}^{p_{m_{\max}}}, \; 1 \leq m \leq M, \tag{89}$$

$$\lambda^{(n+1)} = \lambda^{(n)} - \kappa_2\Delta\lambda^{(n)}, \tag{90}$$

$$w^{(n+1)} = w^{(n)} - \kappa_3\Delta w^{(n)}, \tag{91}$$

$$\mu^{(n+1)} = \mu^{(n)} - \kappa_4\Delta\mu^{(n)}, \tag{92}$$

where again the step sizes are chosen according to (42). The iterative procedure is repeated until $\left\|\mathbf{p}^{(n+1)} - \mathbf{p}^{(n)}\right\|$ becomes smaller than a preset threshold.

*2) ADMM-based method:* Similar to the multi-target case, we reformulate the problem $\mathbb{P}6$ as

$$\begin{array}{ll} \min\limits_{\mathbf{p},\mathbf{z}} & \mathbf{1}^{\mathrm{T}}\mathbf{p}, \\ \text{s.t.} & \eta\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{p} - \mathbf{b}^{\mathrm{T}}\mathbf{p} = 0, \\ & \mathbf{z} = \mathbf{p}, \\ & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \tag{93}$$

Then, by introducing an augmented Lagrangian, we have

$$\begin{array}{ll} \max\limits_{\mathbf{e},\mu}\min\limits_{\mathbf{p},\mathbf{z}} & \mathbf{1}^{\mathrm{T}}\mathbf{p} + \frac{\rho_0}{2}\|\mathbf{p} - \mathbf{z} + \mathbf{e}\|^2 \\ & + \frac{\rho_1}{2}\left\|\eta\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{p} - \mathbf{b}^{\mathrm{T}}\mathbf{p} + \mu\right\|^2, \\ \text{s.t.} & p_{m_{\min}} \leq p_m \leq p_{m_{\max}}, \; 1 \leq m \leq M. \end{array} \tag{94}$$

With the initialization of $\mathbf{p}^{(0)} = \mathbf{z}^{(0)} = \mathbf{p}_{equ}$, $\mathbf{e}^{(0)} = \mathbf{1}$, $\mu^{(0)} = 1$, and $\rho_0^{(0)}$ and $\rho_1^{(0)}$ set to a large positive number, each iteration involves the following steps.

TABLE I
COMPLEXITY PER ITERATION OF THE OCD-BASED ALGORITHM

| Operation | Flops per iteration |
|---|---|
| Update $\mathbf{p}$ | $3KM^2 + (3K+5)\,M + 3K$ |
| Update $\lambda$ | $2KM + 2K + 2,$ |
| Update $\mathbf{w}$ | $2KM^2 + 3KM + 5K,$ |
| Update $\boldsymbol{\mu}$ | $2KM^2 + KM + 4K$ |
| Total | $7KM^2 + (9K+5)M + 14K + 2$ |

TABLE II
COMPLEXITY PER ITERATION OF THE ADMM-BASED ALGORITHM

| Operation | Flops per iteration |
|---|---|
| Update $\mathbf{p}$ | $M^3 + (5K+7)\,M^2 + (4K+8)\,M + 3K + 5$ |
| Update $\mathbf{w}$ | $4KM^2 + (2K^2+4K)M + K^2 + 14K$ |
| Update $\mathbf{z}$ | $M^3 + (7K+2)\,M^2 + (2K+3)\,M + 4K$ |
| Update $\mathbf{e}$ | $2M$ |
| Update $\mu$ | $2KM + 2K + 1$ |
| Update $\boldsymbol{\gamma}$ | $2KM^2 + KM + 3K$ |
| Total | $2M^3 + (18K+9)\,M^2 + \left(2K^2 + 13K + 13\right)M$ |
| | $+K^2 + 26K + 6$ |

- *Step 1: Update* $\mathbf{p}$. Isolating all the terms involving $\mathbf{p}$, the optimization is a constrained convex problem, leading to

$$\begin{aligned}\bar{\mathbf{p}}^{(n+1)} &= \Big(\rho_0^{(n)}\mathbf{I}_M + \rho_1^{(n)}\left(\eta\mathbf{A}^{\mathrm{T}}\mathbf{z}^{(n)} - \mathbf{b}\right)\\ &\quad\times\Big(\eta\big(\mathbf{z}^{(n)}\big)^{\mathrm{T}}\mathbf{A} - \mathbf{b}^{\mathrm{T}}\Big)\Big)^{-1}\Big(-\mathbf{1} + \rho_0^{(n)}\big(\mathbf{z}^{(n)} - \mathbf{e}^{(n)}\big)\\ &\quad - \rho_1^{(n)}\mu^{(n)}\left(\eta\mathbf{A}^{\mathrm{T}}\mathbf{z}^{(n)} - \mathbf{b}\right)\Big), \qquad (95)\end{aligned}$$

$$p_m^{(n+1)} = \left[\bar{p}_m^{(n+1)}\right]_{p_{m\,\min}}^{p_{m\,\max}},\ 1 \le m \le M. \qquad (96)$$

- *Step 2: Update* $\mathbf{z}$. Isolating all the terms involving $\mathbf{z}$, the problem is an unconstrained convex problem, leading to

$$\begin{aligned}\mathbf{z}^{(n+1)} = &\left(\rho_0^{(n)}\mathbf{I}_M + \rho_1^{(n)}\left(\eta\mathbf{A}\mathbf{p}^{(n+1)}\right)\left(\eta\mathbf{A}\mathbf{p}^{(n+1)}\right)^{\mathrm{T}}\right)^{-1}\\ &\times\Big(\rho_0^{(n)}\left(\mathbf{p}^{(n+1)} + \mathbf{e}^{(n)}\right)\\ &+ \rho_1^{(n)}\eta\mathbf{A}\mathbf{p}^{(n+1)}\left(\mathbf{b}^{\mathrm{T}}\mathbf{p}^{(n+1)} - \mu^{(n)}\right)\Big). \qquad (97)\end{aligned}$$

- *Step 3: Update* $\mathbf{e}$ *and* $\mu$. The dual variables are updated according to

$$\mu^{(n+1)} = \mu^{(n)} + \eta\big(\mathbf{z}^{(n+1)}\big)^{\mathrm{T}}\mathbf{A}\mathbf{p}^{(n+1)} - \mathbf{b}^{\mathrm{T}}\mathbf{p}^{(n+1)}, \quad (98)$$

$$\mathbf{e}^{(n+1)} = \mathbf{e}^{(n)} + \mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)}. \qquad (99)$$

- *Step 4: Update the $\rho_0$ and $\rho_1$ at the first a few iterations*. By defining the primal and dual residuals $r_0^{(n+1)}$ and $s_0^{(n+1)}$ as

$$r_0^{(n+1)} = \left\|\mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)}\right\|, \qquad (100)$$

$$s_0^{(n+1)} = \left\|\rho_0^{(n)}\left(\mathbf{z}^{(n)} - \mathbf{z}^{(n+1)}\right)\right\|, \qquad (101)$$

the updated $\rho_0^{(n+1)}$ is given by (75), and the dual variable $\mathbf{e}^{(n+1)}$ is rescaled according to (76). Similarly, define the primal and dual residuals $r_1^{(n+1)}$ and $s_1^{(n+1)}$ as

$$r_1^{(n+1)} = \left|\eta\big(\mathbf{z}^{(n+1)}\big)^{\mathrm{T}}\mathbf{A}\mathbf{p}^{(n+1)} - \mathbf{b}^{\mathrm{T}}\mathbf{p}^{(n+1)}\right|, \qquad (102)$$

$$\begin{aligned}s_1^{(n+1)} &= \Big\|\mu^{(n+1)}\rho_1^{(n)}\eta\mathbf{A}^{\mathrm{T}}\left(\mathbf{z}^{(n)} - \mathbf{z}^{(n+1)}\right) + \rho_1^{(n)}\eta\\ &\quad\times\left(\eta\mathbf{A}^{\mathrm{T}}\mathbf{z}^{(n)} - \mathbf{b}\right)\left(\mathbf{z}^{(n)} - \mathbf{z}^{(n+1)}\right)^{\mathrm{T}}\mathbf{A}\mathbf{p}^{(n+1)}\Big\|. \qquad (103)\end{aligned}$$

The updated $\rho_1^{(n+1)}$ is given by (77), and the rescaled dual variable $\mu^{(n+1)}$ is given by (78).

*3) A closed-form approximate solution:* An equivalent Lagrangian associated with the problem $\mathbb{P}5$ is $L(\mathbf{p},\lambda) = \mathbf{1}^{\mathrm{T}}\mathbf{p} + \lambda\left(\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - \mathbf{b}^{\mathrm{T}}\mathbf{p}\right)$, whose KKT conditions are

$$\mathbf{1} + \lambda\big(\eta\big(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\big)\mathbf{p} - \mathbf{b}\big) = \mathbf{0}, \qquad (104)$$

$$\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p} - \mathbf{b}^{\mathrm{T}}\mathbf{p} = \mathbf{0}. \qquad (105)$$

The authors of [12] obtained the closed-form optimal solution $\lambda^\star$ and $\mathbf{p}^\star$ by jointly solving the two equations (104) and (105). In particular, they calculated $\bar{\mathbf{p}}^\star$ from (104) as

$$\bar{\mathbf{p}}^\star = \frac{\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)^{-1}}{\eta}\left(\mathbf{b} - \frac{1}{\lambda^\star}\mathbf{1}\right), \qquad (106)$$

and then obtained $\mathbf{p}^\star$ by imposing the power constraints

$$p_m^\star = \left[\bar{p}_m^\star\right]_{p_{m\,\min}}^{p_{m\,\max}},\ 1 \le m \le M. \qquad (107)$$

Unfortunately, this closed-form 'optimal' solution is generally invalid because in general $\mathbf{A} + \mathbf{A}^{\mathrm{T}}$ is not invertible.

*Lemma 2:* The rank of $\mathbf{A} + \mathbf{A}^{\mathrm{T}}$ is no more than 3.

*Proof:*

$$\begin{aligned}\operatorname{rank}\big(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\big) &\le \operatorname{rank}\Big(\mathbf{a}_{1,1}\left(\mathbf{a}_{2,2}\right)^{\mathrm{T}} - \mathbf{a}_{1,2}\left(\mathbf{a}_{2,1}\right)^{\mathrm{T}}\\ &\quad + \mathbf{a}_{2,2}\left(\mathbf{a}_{1,1}\right)^{\mathrm{T}} - \mathbf{a}_{2,1}\left(\mathbf{a}_{1,2}\right)^{\mathrm{T}}\Big)\\ &\le \operatorname{rank}\Big(\mathbf{a}_{1,1}\left(\mathbf{a}_{2,2}\right)^{\mathrm{T}}\Big) + \operatorname{rank}\Big(\mathbf{a}_{1,2}\left(\mathbf{a}_{2,1}\right)^{\mathrm{T}}\Big)\\ &\quad + \operatorname{rank}\Big(\mathbf{a}_{2,2}\left(\mathbf{a}_{1,1}\right)^{\mathrm{T}}\Big) \le 3.\end{aligned}$$

The second inequality is due to the fact that $\mathbf{a}_{1,2} = \mathbf{a}_{2,1}$.

Clearly, for any system with more than 3 transmit radars, the solution of (106) is invalid, and the minimum eigenvalue $\xi_{\min}$ of $\mathbf{A} + \mathbf{A}^{\mathrm{T}}$ is negative. We propose an approximate closed-form solution by replacing the invalid $\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)^{-1}$ in (106) by the valid regularized form

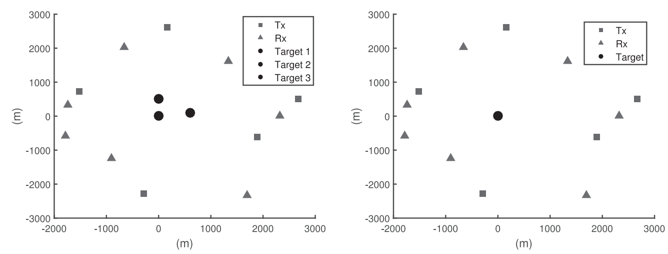
$$\mathbf{B} = \left(\mathbf{A} + \mathbf{A}^{\mathrm{T}} + \left(|\xi_{\min}| + \epsilon\right)\mathbf{I}_M\right)^{-1}, \qquad (108)$$

where $\epsilon$ is a small positive number, such as, $\epsilon = 0.0001$. Thus the 'unconstrained' power allocation is given as

$$\bar{\mathbf{p}}^{\circledast} = \frac{\mathbf{B}}{\eta}\left(\mathbf{b} - \frac{1}{\lambda^{\circledast}}\mathbf{1}\right), \qquad (109)$$

TABLE III
COMPLEXITY PER ITERATION OF THE ALGORITHM GIVEN IN [13], WHERE $n_{\text{in}}$ IS THE AVERAGE NUMBER OF INNER ITERATIONS IN INNER OPTIMIZATION PROCEDURE

| Operation | Flops per inner iteration |
|---|---|
| Update the parameters of inner QCLP problem | $\left(5M^2+2M+1\right)/n_{\text{in}}$ |
| Solve the inner QCLP problem | $M^3+(4K+3)\,M^2+(6K+10)M-K$ |
| Total | $M^3+\left(4K+3+\frac{5}{n_{\text{in}}}\right)M^2+\left(6K+10+\frac{2}{n_{\text{in}}}\right)M-K$ |




(a) (b)

Fig. 2. Illustration of the MIMO radar system for: (a) three-target application, and (b) single-target application.

where $\lambda^{*}$ is obtained by substituting $\bar{\mathbf{p}}^{*}$ into (105) and taking the positive solution as

$$\lambda^{*}=\frac{-b+\sqrt{b^2-4ac}}{2a}, \tag{110}$$

with

$$\begin{cases} a=\mathbf{b}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}\mathbf{B}\mathbf{b}-\mathbf{b}^{\mathrm{T}}\mathbf{B}\mathbf{b}, \\ b=-2\mathbf{1}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{B}\mathbf{b}^{\mathrm{T}}+2\mathbf{b}^{\mathrm{T}}\mathbf{B}\mathbf{1}, \\ c=\mathbf{1}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}\mathbf{B}\mathbf{1}-\mathbf{1}^{\mathrm{T}}\mathbf{B}\mathbf{1}. \end{cases} \tag{111}$$

The solution $\mathbf{p}^{*}$ is then obtained by projecting $\bar{\mathbf{p}}^{*}$ onto the feasible region. This closed-form solution is inferior to the OCD-based and ADMM-based solutions in terms of its achievable performance, owing to its suboptimal nature.

## IV. CONVERGENCE AND COMPLEXITY ANALYSIS

Recall from Section II and III that our optimization problem $\mathbb{P}1$ of (24) is nonconvex, and both our ADMM and OCD algorithms are based on a Lagrangian function approach. It is widely acknowledged that the zero duality gap cannot be guaranteed for general nonconvex problems. However, Yu and Lui [24] proposed a theorem which guarantees the zero duality gap for the nonconvex problem that meets the 'time-sharing condition'. In Appendix B, we proved that our optimization problem $\mathbb{P}1$ satisfies the time-sharing condition of [24]. Hence, the strong duality holds for $\mathbb{P}1$. We are now ready to prove that both our two algorithms can converge to a local optimal point under some assumptions.

### A. Convergence of the Proposed Algorithms

*1) The ADMM-based algorithm:* We first point out again that since our problem is nonconvex, the ADMM-based algorithm can only guarantee to converge to a local optimal solution. The convergence of the ADMM method is proved for the convex optimization problem in [18], while Magnússon *et al.* [25] extended the convergence results to the nonconvex case. The convergence of our ADMM-based algorithm will be further illustrated in Section V using simulations.

*2) The OCD-based algorithm:* Again, since our optimization problem is nonconvex, the OCD-based algorithm can only find a locally optimal solution. Collect all the primal variables of the Lagrangian (32) together as $\mathbf{y}=\left[\mathbf{p}^{\mathrm{T}}\ \mathbf{w}^{\mathrm{T}}\right]^{\mathrm{T}}$ and denote the cost function and the constraints of $\mathbb{P}3$ respectively by

$$f(\mathbf{y})=\mathbf{1}^{\mathrm{T}}\mathbf{p}, \tag{112}$$

$$g_0(\mathbf{y})=\sum_{k=1}^{K}v_k w_k\mathbf{b}_k^{\mathrm{T}}\mathbf{p}-1, \tag{113}$$

$$g_k(\mathbf{y})=w_k\eta\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}-1,\ 1\le k\le K. \tag{114}$$

According to Theorem 2 in Section 8.2.3 and Lemma 5 in Section 2.1 of [26], to prove the convergence of the OCD algorithm, we have to prove that the second derivatives $\nabla^2 f(\mathbf{y})$ and $\nabla^2 g_k(\mathbf{y})$ for $0\le k\le K$ satisfy the Lipschitz condition in a neighbourhood of the optimal primal point $\mathbf{y}^{\star}$. Note that

$$\nabla^2 f(\mathbf{y})=\mathbf{0}, \tag{115}$$

$$\nabla^2 g_0(\mathbf{y})=\begin{bmatrix} \mathbf{0} & v_1\mathbf{b}_1\cdots v_K\mathbf{b}_K \\ v_1\mathbf{b}_1^{\mathrm{T}} & \\ \vdots & \mathbf{0} \\ v_K\mathbf{b}_K^{\mathrm{T}} & \end{bmatrix}, \tag{116}$$

$$\nabla^2 g_k(\mathbf{y})=\eta\begin{bmatrix} w_k\left(\mathbf{A}_k+\mathbf{A}_k^{\mathrm{T}}\right)\ \mathbf{0} & \left(\mathbf{A}_k+\mathbf{A}_k^{\mathrm{T}}\right)\mathbf{p}\ \mathbf{0} \\ \mathbf{0} & \\ \left(\mathbf{A}_k+\mathbf{A}_k^{\mathrm{T}}\right)\mathbf{p}^{\mathrm{T}} & \mathbf{0} \\ \mathbf{0} & \end{bmatrix},$$
$$1\le k\le K. \tag{117}$$

Since $\nabla^2 f(\mathbf{y})$ and $\nabla^2 g_0(\mathbf{y})$ are constants, they satisfy the required Lipschitz condition. For $\mathbf{p}_{\min}\le\mathbf{p}\le\mathbf{p}_{\max}$, all the elements in the matrix $\nabla^2 g_k(\mathbf{y})$, where $1\le k\le K$, are finite. Therefore, it is easy to find a constant $\varsigma$ satisfying

$$\left\|\nabla^2 g_k(\mathbf{y}_1)-\nabla^2 g_k(\mathbf{y}_2)\right\|\le\varsigma\left\|\mathbf{y}_1-\mathbf{y}_2\right\|. \tag{118}$$

Thus $\nabla^2 g_k(\mathbf{y})$ satisfies the required Lipschitz condition too.

According to [26], under the assumption that the Hessian matrix of the Lagrangian (32) with respect to $\mathbf{y}$ at the minimum primal point $\mathbf{y}^{\star}=\left(\mathbf{p}^{\star},\mathbf{w}^{\star}\right)$ is positive definite, the Hessian matrix of the Lagrangian (32) with respect to the primal and dual variables is negative definite at the optimal point $\left(\mathbf{p}^{\star},\mathbf{w}^{\star},\lambda^{\star},\boldsymbol{\mu}^{\star}\right)$. Then there exists a positive number $\overline{\kappa}=\min_i -\Re\left[\bar{\xi}_i\right]\left|\bar{\xi}_i\right|^{-2}$,

TABLE IV
SYSTEM PARAMETERS

| | Parameters | Values |
|---|---|---|
| | Effective bandwidth $\beta_m$ | 200 MHz |
| | Transmit power upper bound $p_{m_{\max}}$ | 300 watts |
| | Transmit power lower bound $p_{m_{\min}}$ | 1 watts |
| | Transmit radars' positions | $(2665, 508)$, $(165, 2617)$, $(-1520, 715)$, $(-287, -2270)$, $(1892, -615)$ |
| | Receive radars' positions | $(2320, 0)$, $(1338, 1617)$, $(-656, 2019)$, $(-1740, 332)$, $(-1791, -582)$, $(-900, -1238)$, $(1696, -2334)$ |
| | Path loss $\kappa_{m,n}^{(k}$ | $\kappa_{m,n}^{(k} = \frac{1}{10\left(R_{m,k}^{tx}\right)^2\left(R_{n,k}^{rx}\right)^2}$ |
| Three targets | Targets' positions $(x_k, y_k)$ | $(0, 0)$, $(0, 500)$, $(600, 100)$ |
| | RCS model for target 1,$\left\|\mathbf{h}^{(1}\right\|$ | $\begin{matrix} 0.75 & 0.4 & 0.45 & 0.55 & 0.3 & 0.2 & 0.25 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.1 & 0.05 & 0.01 & 0.12 & 0.09 & 0.2 & 0.19 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$ |
| | RCS model for target 2,$\left\|\mathbf{h}^{(2}\right\|$ | $\begin{matrix} 0.1 & 0.05 & 0.01 & 0.12 & 0.09 & 0.2 & 0.19 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.75 & 0.4 & 0.45 & 0.55 & 0.3 & 0.2 & 0.25 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$ |
| | RCS model for target 3,$\left\|\mathbf{h}^{(3}\right\|$ | $\begin{matrix} 0.1 & 0.05 & 0.01 & 0.12 & 0.09 & 0.2 & 0.19 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.75 & 0.4 & 0.45 & 0.55 & 0.3 & 0.2 & 0.25 \end{matrix}$ |
| Single target | Target's position $(x, y)$ | $(0, 0)$ |
| | RCS model for target,$\left\|\mathbf{h}\right\|$ | $\begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.01 & 0.05 & 0.01 & 0.022 & 0.092 & 0.092 & 0.092 \\ 0.45 & 0.35 & 0.48 & 0.32 & 0.49 & 0.49 & 0.49 \\ 0.22 & 0.55 & 0.55 & 0.48 & 0.57 & 0.57 & 0.57 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$ |

where $\bar{\xi}_i$ are the eigenvalues of the Hessian matrix of the Lagrangian (32) with respect to the primal and dual variables at $\left(\mathbf{p}^\star, \mathbf{w}^\star, \lambda^\star, \boldsymbol{\mu}^\star\right)$. Consequently, as long as the maximum of the four step sizes $\kappa_{\max} = \max_{1 \le i \le 4} \kappa_i$ satisfies the condition of $\kappa_{\max} \le \overline{\kappa}$, our scheme (37)–(40) will converge to the locally optimal point $\left(\mathbf{p}^\star, \mathbf{w}^\star, \lambda^\star, \boldsymbol{\mu}^\star\right)$ when starting from a neighbourhood of $\left(\mathbf{p}^\star, \mathbf{w}^\star, \lambda^\star, \boldsymbol{\mu}^\star\right)$, according to [26]. In practice, $\overline{\kappa}$ is unknown. It is advisable to choose sufficiently small step sizes $\kappa_i$, $1 \le i \le 4$, in order to ensure the convergence of the OCD scheme.

*Remark 4*: A positive-definite Hessian matrix of the Lagrangian (32) with respect to $\mathbf{y}$ at $\mathbf{y}^\star$ is a sufficient condition for the convergence of the OCD scheme. If this Hessian matrix is semi-positive definite, we cannot prove the convergence of the OCD scheme based on the result of [26]. By adopting an exponentially decaying step size $\kappa_{\max} \propto e^{-\alpha n}$, we ensure that our OCD algorithm works well in any situation.

## B. Complexity of Proposed Algorithms and Algorithm of [13]

The complexity of our OCD and ADMM algorithms are summarized in Tables I and II, respectively. For the ADMM-based algorithm, since the penalty parameters are only updated in the first few iterations, the complexity associated with this part of operation is omitted. Additionally, we assume that Gauss-Jordan elimination is used for matrix inversion and, therefore, the number of flops required by inverting an $M \times M$ matrix is $M^3 + M^2 + M$. For the OCD-based algorithm, the complexity of computing the four step sizes is negligible and therefore it is also omitted. Clearly, the complexity of the ADMM-based algorithm is on the order of $M^3$ per iteration, which is denoted by $\mathsf{O}\left(M^3\right)$, while the complexity of the OCD-based algorithm is on the order of $\mathsf{O}\left(M^2\right)$ per iteration. It will be shown by our simulation results that the convergence speed of the ADMM algorithm is at least one order of magnitude faster than that of the OCD algorithm. Therefore, despite its higher per-iteration complexity, the ADMM algorithm actually imposes a lower total complexity, compared to the OCD algorithm.

The benchmark scheme of [13] invokes two iterative loops for solving the optimization problem (25). Specifically, at each outer iteration, the parameters of the inner quadratic constrained linear programming (QCLP) problem are updated, and the QCLP problem is then solved iteratively in the inner iterative loop. We assume that the interior-point method is used for solving this inner QCLP, which requires $n_{\text{in}}$ iterations on average. Based on the above discussions, the complexity of the algorithm of [13] is summarized in Table III, where it is seen that the complexity per inner iteration is on the order of $\mathsf{O}\left(M^3\right)$. Thus the complexity of our ADMM-based algorithm is only marginally higher than that of the algorithm in [13], because they are both on the order of $\mathsf{O}\left(M^3\right)$ per iteration. The algorithm of [13] requires a total of $n_{\text{ou}} n_{\text{in}}$ iterations to converge, where $n_{\text{ou}}$ is the number of iterations for the outer iterative loop. As it will be shown in the simulation results, the number of iterations required for the ADMM-based algorithm to converge is very close to the total number of iterations $n_{\text{ou}} n_{\text{in}}$ required by the algorithm of [13].

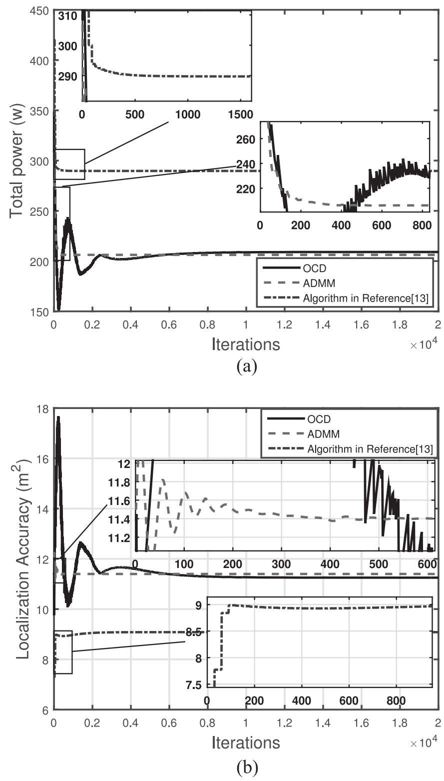



Fig. 3. Convergence performance of three algorithms, in terms of (a) total power consumption, and (b) aggregate localization accuracy, for the three-target case with $v_1 = 1$, $v_2 = 2$ and $v_3 = 1$.

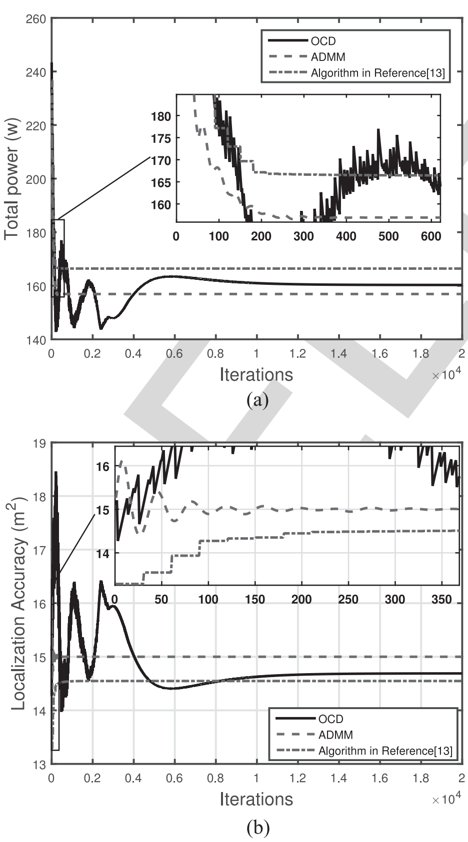



Fig. 4. Convergence performance of three algorithms, in terms of (a) total power consumption, and (b) aggregate localization accuracy, for the three-target case with $v_1 = v_2 = v_3 = 1$.

In this sense, both algorithms require a similar total complexity for solving their associated optimization problems. Although our OCD-based algorithm enjoys a much lower complexity per iteration than the algorithm of [13], it imposes a higher total complexity.

## V. Simulation Results

Let us now evaluate the performance of the proposed algorithms using a MIMO radar system having $M = 5$ transmit radars and $N = 7$ receive radars. The algorithm of [13] is used as the benchmark. Fig. 2 depicts both the triple-target and single-target cases considered. The system parameters of both the triple-target and single-target cases are listed in Table IV. The localization accuracy threshold $\eta$ is set to 15 m$^2$ for the triple-target case and 10 m$^2$ for the single-target case. The exponential decaying factor is empirically chosen to be $\alpha = 0.0005$ for the four step sizes of the OCD algorithm.

### A. Triple-Target Case

We consider the two sets of the importance weightings for the three targets given by: i) $v_1 = 1$, $v_2 = 2$ and $v_3 = 1$, and ii) $v_1 = v_2 = v_3 = 1$. For the sake of a fair comparison to the algorithm of [13], the effects of these weightings have to be taken into consideration, and the target estimation error thresholds for the three constraints of the optimization problem (25) are suitably scaled as

$$\frac{\mathbf{b}_1^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_1\mathbf{p}} \le \bar{\eta}_1, \ \frac{\mathbf{b}_2^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_2\mathbf{p}} \le \bar{\eta}_2, \ \frac{\mathbf{b}_3^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_3\mathbf{p}} \le \bar{\eta}_3,$$

with $\bar{\eta}_1 = \frac{1}{3v_1}\eta$, $\bar{\eta}_2 = \frac{1}{3v_2}\eta$ and $\bar{\eta}_3 = \frac{1}{3v_3}\eta$. For our ADMM algorithm, the initial values of the dual variables are set to $\mathbf{e}^{(0)} = [1\,1\,1\,1\,1]^{\mathrm{T}}$, $\mu^{(0)} = 1$ and $\gamma_k^{(0)} = 1$ for $1 \le k \le 3$, while all the initial penalty parameters are set to 500. For our OCD algorithm, the initial values of the dual variables are set to $\lambda^{(0)} = 1$ and $\mu_k^{(0)} = 1$ for $1 \le k \le 3$. Additionally, the four constants in the four step sizes of the OCD algorithm are set to $c_1 = 0.3$, $c_2 = 1.0$, $c_3 = 1.5$ and $c_4 = 1.1$ for the senario i), while they are set to $c_1 = 0.3$, $c_2 = 0.9$, $c_3 = 1.5$ and $c_4 = 1.1$ for the senario ii). These parameters were found empirically to be appropriate for the corresponding application scenarios. For the algorithm of [13], we use the CVX software to solve its inner QCLP problem. In our simulations, we observe that the CVX converges within 25 to 35 iterations. Therefore, we will assume that the average number of inner iterations for the algorithm of [13] is $n_{\mathrm{in}} = 30$. Q1

Fig. 3 compares the total power allocations $\mathbf{p}$ and the aggregate localization accuracy results of $\sum_{k=1}^{3} \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}}{\mathbf{p}^{\mathrm{T}}\mathbf{A}_k\mathbf{p}}$ obtained by the three algorithms for the senario i), while Fig. 4 depicts the results for the senario ii). It can be seen that the number of iterations required by the ADMM-based algorithm to converge is similar to the total number of iterations $n_{\mathrm{ou}}n_{\mathrm{in}}$ required by the algorithm of [13], while the convergence speed of the OCD-based algorithm is considerably slower than that of the other two algorithms. As expected, our algorithms outperform the algorithm of [13] in terms of its total power consumption, albeit at the expense of some degradation in localization accuracy.

TABLE V
PERFORMANCE COMPARISON OF THREE ALGORITHMS FOR THE THREE-TARGET CASE

| | ii) $v_1 = v_2 = v_3 = 1$ | | | i) $v_1 = 1,\ v_2 = 2,\ v_3 = 1$ | | |
|---|---|---|---|---|---|---|
| | ADMM | OCD | [13] | ADMM | OCD | [13] |
| Radar 1: Power (watts) | 1 | 1 | 1 | 1 | 1 | 1 |
| Radar 2: Power (watts) | 95.8 | 93.3 | 102 | 119.6 | 117.9 | 75.8 |
| Radar 3: Power (watts) | 58.2 | 64.0 | 40.3 | 83.5 | 88.1 | 170.4 |
| Radar 4: Power (watts) | 1 | 1 | 1 | 1 | 1 | 1 |
| Radar 5: Power (watts) | 1 | 1 | 22.2 | 1 | 1 | 41.5 |
| Total Power (watts) | 157 | 160.3 | 166.5 | 206.1 | 209.0 | 289.7 |
| Target 1: Localization Accuracy ($m^2$) | 5.4 | 5.3 | 5 | 4.1 | 4.1 | 3.1 |
| Target 2: Localization Accuracy ($m^2$) | 4.8 | 4.7 | 4.5 | 3.6 | 3.5 | 2.5 |
| Target 3: Localization Accuracy ($m^2$) | 4.8 | 4.7 | 5 | 3.7 | 3.6 | 3.5 |
| Aggregate Localization Accuracy ($m^2$) | 15 | 14.7 | 14.5 | 11.4 | 11.3 | 9.1 |
| Total Power Saving | 5.7% | 3.7% | - | 28.9% | 27.9% | - |
| Degradation in Aggregate Localization Accuracy | 3.4% | 1.4% | - | 25.3% | 27.9% | - |
| Average Total Power Saving | 10.0% | 10.5% | - | 20.0 % | 25.6 % | - |
| Average Degradation in Aggregate Localization Accuracy | 8.6% | 8.9% | - | 27.2% | 30.0% | - |

The average results are obtained over 1000 random simulation experiments.

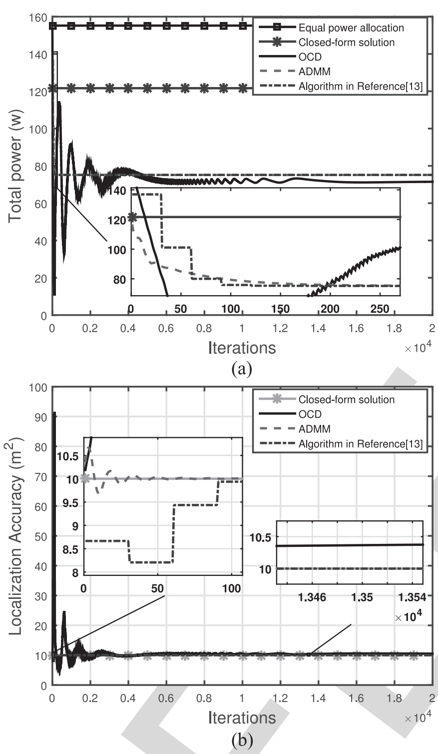



Fig. 5. Convergence performance of three algorithms, in terms of (a) total power consumption, and (b) aggregate localization accuracy, in comparison with the EPA and the closed-form solution, for the single-target case.

Table V details how our algorithms trade the localization accuracy against the transmit power, in comparison to the algorithm of [13]. Specifically, for the senario of i), our ADMM algorithm achieves 28.9% power saving at the cost of 25.3% degradation in aggregate localization accuracy, while our OCD algorithm trades 27.9% power saving against 27.9% degradation in localization accuracy. For the equal weighting senario of ii), the savings in power achieved by our two algorithms are considerably smaller but the losses in localization accuracy are also significantly smaller, compared with the senario i). To obtain statistically relevant comparison, we carry out 1000 simulations by randomly locating all the transmit radars and receive radars at the radius $R = 3000(0.5 + \varepsilon_x)$ m with the angular rotations of $\theta = 2\pi\varepsilon_y$, where $\varepsilon_x$ and $\varepsilon_y$ are uniformly distributed in $[0,\ 1.0]$. The average power saving and degradation in localization accuracy over the 1000 random experiments are listed in the last two rows of Table V.

### B. Single-Target Case

The four constants in the four step sizes of the OCD algorithm are set to $c_1 = c_2 = 1.0$ and $c_3 = c_4 = 0.3$, which is empirically found to be appropriate for this application senario. Fig. 5 characterizes the performance of our ADMM-based and OCD-based algorithms as well as the algorithm of [13]. As expected, all the three algorithms attain the same performance, both in terms of total power allocated and localization accuracy, since the underlying optimization problems are identical in the single-target case. In terms of convergence speed, our ADMM-based algorithm outperforms the algorithm of [13], while the OCD-based algorithm is considerably slower than the algorithm of [13]. In Fig. 5 (a), we also characterize the equal-power allocation (EPA) scheme and the closed-form solution of SubSection III-B3. It can be seen that our closed-form solution performs significantly better than the EPA scheme, but it is inferior to the other three iterative algorithms because the suboptimal nature of this closed-form solution.

## VI. CONCLUSION

The target localization problem of distributed MIMO radar systems has been investigated, which minimizes the power of the transmit radars, while meeting a required localization accuracy. We have proposed the OCD-based and ADMM-based iterative algorithms to solve this nonconvex optimization problem. Both the algorithms are capable of converging to a local optimum. The OCD algorithm imposes a much lower computational complexity per iteration, while the ADMM algorithm achieves a much faster convergence. For the multi-target senario, we have shown how our proposed approach trades the power saving with some degradation in localization accuracy, compared with that of state-of-the-art scheme [13]. We have also demonstrated that our ADMM-based algorithm and the existing state-of-the-art scheme have similar computational complexity and convergence speed. For the single-target senario, we have confirmed that our algorithms and the benchmark attain the same performance in terms of both power consumption and localization accuracy, because the underlying optimization problems become identical.

## Appendix

### A. Derivation of Updating Formulae for Penalty Parameters

The optimal solution to the $\mathbb{P}4$ of (45) should be primal and dual feasible, that is,

$$\mathbf{p}^{(n+1)} - \mathbf{z}^{(n+1)} = \mathbf{0}, \tag{119}$$

$$\sum_{k=1}^{K} w_k^{(n+1)} v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p}^{(n+1)} - 1 = 0, \tag{120}$$

$$w_k \left(\mathbf{z}^{(n+1)}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{p}^{(n+1)} \eta - 1 = 0,\ 1 \le k \le K, \tag{121}$$

$$\frac{\partial L'\left(\mathbf{p}, \mathbf{z}^{(n+1)}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{p}} = \mathbf{0}, \tag{122}$$

$$\frac{\partial L'\left(\mathbf{p}^{(n+1)}, \mathbf{z}^{(n+1)}, \mathbf{w}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{w}} = \mathbf{0}, \tag{123}$$

$$\frac{\partial L'\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{z}} = \mathbf{0}, \tag{124}$$

where $L'\left(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_0, d_1, \mathbf{d}_2\right)$ is the Lagrangian of (45), which can be separated into three parts

$$L'\left(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_0, d_1, \mathbf{d}_2\right) = \underbrace{\mathbf{1}^{\mathrm{T}}\mathbf{p} + \mathbf{d}_0^{\mathrm{T}}(\mathbf{p} - \mathbf{z})}_{L'_0(\mathbf{p}, \mathbf{z}, \mathbf{d}_0)} + \underbrace{d_1 \left(\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p} - 1\right)}_{L'_1(\mathbf{p}, \mathbf{w}, d_1)} + \underbrace{\sum_{k=1}^{K} d_{2,k} \left(w_k \mathbf{z}^{\mathrm{T}} \mathbf{A}_k \mathbf{p} \eta - 1\right)}_{L'_2(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_2)}. \tag{125}$$

However, the ADMM-based algorithm uses the augmented Lagrangian of

$$\begin{aligned} L(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_0, d_1, \mathbf{d}_2) &= \underbrace{\mathbf{1}^{\mathrm{T}}\mathbf{p} + \frac{\rho_0}{2}\|\mathbf{p} - \mathbf{z}\|^2 + \mathbf{d}_0^{\mathrm{T}}(\mathbf{p} - \mathbf{z})}_{L_0(\mathbf{p}, \mathbf{z}, \mathbf{d}_0)} \\ &+ \underbrace{\frac{\rho_1}{2}\left|\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p} - 1\right|^2 + d_1 \left(\sum_{k=1}^{K} w_k v_k \mathbf{b}_k^{\mathrm{T}} \mathbf{p} - 1\right)}_{L_1(\mathbf{p}, \mathbf{w}, d_1)} \\ &+ \underbrace{\sum_{k=1}^{K} \frac{\rho_{2,k}}{2} \left|w_k \mathbf{z}^{\mathrm{T}} \mathbf{A}_k \mathbf{p} \eta - 1\right|^2 + \sum_{k=1}^{K} d_{2,k} \left(w_k \mathbf{z}^{\mathrm{T}} \mathbf{A}_k \mathbf{p} \eta - 1\right)}_{L_2(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_2)}, \end{aligned} \tag{126}$$

which can be divided into three parts, and all the primal and dual variables are updated one by one. Thus, in every iteration, there exist primal and dual residuals.

Specifically, in the $(n+1)$th iteration, the primal residuals are given by $r_0^{(n+1)}$ of (65), $r_1^{(n+1)}$ of (66), and $r_{2,k}^{(n+1)}$ for $1 \le k \le K$ of (67), while the dual residuals are defined via

$$dr = \sqrt{\|\mathbf{dr}_0\|^2 + \|\mathbf{dr}_1\|^2 + \|\mathbf{dr}_2\|^2}, \tag{127}$$

with

$$\begin{aligned} \mathbf{dr}_0 = &\frac{\partial L\left(\mathbf{p}, \mathbf{z}^{(n)}, \mathbf{w}^{(n)}, \mathbf{d}_0^{(n)}, d_1^{(n)}, \mathbf{d}_2^{(n)}\right)}{\partial \mathbf{p}} \\ &- \frac{\partial L'\left(\mathbf{p}, \mathbf{z}^{(n+1)}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{p}}, \end{aligned} \tag{128}$$

$$\begin{aligned} \mathbf{dr}_1 = &\frac{\partial L\left(\mathbf{p}^{(n+1)}, \mathbf{z}^{(n)}, \mathbf{w}, \mathbf{d}_0^{(n)}, d_1^{(n)}, \mathbf{d}_2^{(n)}\right)}{\partial \mathbf{w}} \\ &- \frac{\partial L'\left(\mathbf{p}^{(n+1)}, \mathbf{z}^{(n+1)}, \mathbf{w}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{w}}, \end{aligned} \tag{129}$$

$$\begin{aligned} \mathbf{dr}_2 = &\frac{\partial L\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n)}, d_1^{(n)}, \mathbf{d}_2^{(n)}\right)}{\partial \mathbf{z}} \\ &- \frac{\partial L'\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{w}^{(n+1)}, \mathbf{d}_0^{(n+1)}, d_1^{(n+1)}, \mathbf{d}_2^{(n+1)}\right)}{\partial \mathbf{z}}. \end{aligned} \tag{130}$$

It can be seen that the primal residuals $r_0^{(n+1)}$, $r_1^{(n+1)}$ and $r_{2,k}^{(n+)}$ for $1 \le k \le K$ are related to $L_0(\mathbf{p}, \mathbf{z}, \mathbf{d}_0)$, $L_1(\mathbf{p}, \mathbf{w}, d_1)$ and $L_2(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_2)$, respectively. Therefore, we will similarly 'separate' the dual residuals into $s_0^{(n+1)}$, $s_1^{(n+1)}$ and $s_{2,k}^{(n+1)}$ for $1 \le k \le K$, corresponding to $L_0(\mathbf{p}, \mathbf{z}, \mathbf{d}_0)$, $L_1(\mathbf{p}, \mathbf{w}, d_1)$ and $L_2(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_2)$, respectively.

In order to analyze the updating formula (75) for the penalty parameter $\rho_0$, we have to calculate $s_0^{(n+1)}$ as follows

$$\begin{aligned} s_0^{(n+1)} = \Bigg( &\left\|\frac{\partial L_0\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{d}_0^{(n)}\right)}{\partial \mathbf{z}} - \frac{\partial L'_0\left(\mathbf{p}^{(n+1)}, \mathbf{z}, \mathbf{d}_0^{(n+1)}\right)}{\partial \mathbf{z}}\right\|^2 \\ &+ \left\|\frac{\partial L_0\left(\mathbf{p}, \mathbf{z}^{(n)}, \mathbf{d}_0^{(n)}\right)}{\partial \mathbf{p}} - \frac{\partial L'_0\left(\mathbf{p}, \mathbf{z}^{(n+1)}, \mathbf{d}_0^{(n+1)}\right)}{\partial \mathbf{p}}\right\|^2 \Bigg)^{\frac{1}{2}}. \end{aligned} \tag{131}$$

By evaluating the required four partial derivatives and plugging them into (131), we arrive at the dual residual $s_0^{(n+1)}$ of (68). Note that a large value for $\rho_0$ adds a large penalty on the violation of primal feasibility and, therefore, a large $\rho_0$ reduces the primal residual $r_0^{(n+1)}$. On the other hand, from the expression (68), it is seen that a small $\rho_0$ reduces the dual residual $s_0^{(n+1)}$. Thus, in order to balance the primal and dual residuals $r_0^{(n+1)}$ and $s_0^{(n+1)}$, the penalty parameter $\rho_0$ is updated according to (75), which is beneficial to convergence.

Similarly, it can be shown that the dual residual $s_1^{(n+1)}$ related to $L_1(\mathbf{p}, \mathbf{w}, d_1)$ is given by (69) and (71), while the dual residuals $s_{2,k}^{(n+1)}$ for $1 \le k \le K$ related to $L_2(\mathbf{p}, \mathbf{w}, \mathbf{z}, \mathbf{d}_2)$ are specified by (70), (72) and (73). Following the same logic of balancing the primal and dual residuals, the updating formulae for the penalty parameters $\rho_1$ and $\rho_{2,k}$ are specified by (76) and (78), respectively.

### *B. Proof of the Time-Sharing Condition for Problem $\mathbb{P}1$*

According to [24], the time-sharing condition for the optimization problem $\mathbb{P}1$ of (24) is as follows. *Time-sharing condition*: Let $\mathbf{p}_1$ and $\mathbf{p}_2$ be the optimal solutions of $\mathbb{P}1$ in conjunction with $\eta = \eta_1$ and $\eta = \eta_2$, respectively. $\mathbb{P}1$ is said to satisfy the time-sharing condition if for any $\eta_1$ and $\eta_2$ and for any $0 \le \xi \le 1$, there always exists a feasible solution $\mathbf{p}_3$ so that $\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_3} \le \xi\eta_1 + (1-\xi)\eta_2$ and $\mathbf{1}^{\mathrm{T}}\mathbf{p}_3 \ge \xi\mathbf{1}^{\mathrm{T}}\mathbf{p}_1 + (1-\xi)\mathbf{1}^{\mathrm{T}}\mathbf{p}_2$.

According to Lemma 1, if we set $\mathbf{p}_3 = \mathbf{p}_{\max}$, then

$$\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_3} \le \eta_1 \text{ and } \sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_3} \le \eta_2.$$

Hence

$$\begin{aligned}\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_3} &= \xi \sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_3} \\ &\quad + (1-\xi)\sum_{k=1}^{K} v_k \frac{\mathbf{b}_k^{\mathrm{T}}\mathbf{p}_3}{\mathbf{p}_3^{\mathrm{T}}\mathbf{A}_k\mathbf{p}_3} \le \xi\eta_1 + (1-\xi)\eta_2, \\ \mathbf{1}^{\mathrm{T}}\mathbf{p}_3 = \xi\mathbf{1}^{\mathrm{T}}\mathbf{p}_3 + (1-\xi)\mathbf{1}^{\mathrm{T}}\mathbf{p}_3 &\ge \xi\mathbf{1}^{\mathrm{T}}\mathbf{p}_1 + (1-\xi)\mathbf{1}^{\mathrm{T}}\mathbf{p}_2.\end{aligned}$$

Therefore, $\mathbb{P}1$ satisfies the time-sharing condition and the dual gap for our nonconvex problem is zero.

## REFERENCES


[1] J. Li and P. Stoica, *MIMO Radar Signal Processing*. Hoboken, NJ, USA: Wiley, 2009.

[2] Y. Yu, A. P. Petropulu, and H. V. Poor, "Measurement matrix design for compressive sensing–based MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5338–5352, Nov. 2011.

[3] S. Gogineni and A. Nehorai, "Target estimation using sparse modeling for distributed MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5315–5325, Nov. 2011.

[4] Y. Yao, A. P. Petrppulu, and H. V. Poor, "CSSF MIMO radar: Compressive-sensing and step-frequency based MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1490–1504, Apr. 2012.

[5] J. Xu, X.-Z. Dai, X.-G. Xia, L.-B. Wang, J. Yu, and Y.-N. Peng, "Optimization of multisite radar system with MIMO radars for target detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 4, pp. 2329–2343, Oct. 2011.

[6] E. Fishler, A. Haimovich, R. S. Blum, L. J. Cimini, D. Chizhik, and R. A. Valenzuela, "Spatial diversity in radars-models and detection performance," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 823–838, Mar. 2006.

[7] Q. He, N. H. Lehmann, R. S. Blum, and A. M. Haimovich, "MIMO-radar moving target detection in homogeneous clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 3, pp. 1290–1301, Jul. 2010.

[8] Q. He, R. S. Blum, H. Godrich, and A. M. Haimovich, "Cramer-Rao bound for target velocity estimation in MIMO radar with widely separated antennas," in *Proc. 42nd Annu. Conf. Inf. Sci. Syst.*, Mar. 19–21, 2008, pp. 123–127.

[9] H. Godrich, A. M. Haimovich, and R. S. Blum, "Target localisation techniques and tools for multiple-input multiple-output radar," *IET Radar, Sonar Navigat.*, vol. 3, no. 4, pp. 314–327, Aug. 2009.

[10] Q. He, R. S. Blum, H. Godrich, and A. M. Haimovich, "Target velocity estimation and antenna placement for MIMO radar with widely separated antennas," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 79–100, Feb. 2010.

[11] H. Godrich, A. M. Haimovich, and R. S. Blum, "Target localization accuracy gain in MIMO radar-based systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2783–2803, Jun. 2010.

[12] H. Godrich, A. P. Petropulu, and H. V. Poor, "Power allocation strategies for target localization in distributed multiple-radar architectures," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3226–3240, Jul. 2011.

[13] N. Garcia, A. M. Haimovich, M. Coulon, and M. Lops, "Resource allocation in MIMO radar with multiple targets for non-coherent localization," *IEEE Trans. Signal Process.*, vol. 62, no. 10, pp. 2656–2666, May 2014.

[14] N. H. Lehmann *et al.*, "Evaluation of transmit diversity in MIMO-radar direction finding," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2215–2225, May 2007.

[15] C. Wei, Q. He, and R. S. Blum, "Cramer-Rao bound for joint location and velocity estimation in multi-target non-coherent MIMO radars," in *Proc. 44th Annu. Conf. Inf. Sci. Syst.*, Mar. 17–19, 2010, pp. 1–6.

[16] Q. He and R. S. Blum, "CramerRao bound for MIMO radar target localization with phase errors," *IEEE Signal Process. Lett.*, vol. 17, no. 1, pp. 83–86, Jan. 2010.

[17] A. J. Conejo, E. Castillo, R. Minguez, and R. Garcia-Bertrand, *Decomposition Techniques in Mathematical Programming: Engineering and Science Applications*. Berlin, Germany: Springer-Verlag, 2006.

[18] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[19] A. Simonetto and G. Leus, "Distributed maximum likelihood sensor network localization," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1424–1437, Mar. 2014.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[21] H. Godrich, A. M. Haimovich, and R. S. Blum, "Cramer Rao bound on target localization estimation in MIMO radar systems," in *Proc. 42nd Annu. Conf. Inf. Sci. Syst.*, Mar. 19–21, 2008, pp. 134–139.

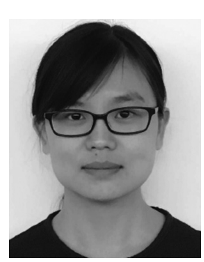
[22] H. Godrich, A. P. Petropulu, and H. V. Poor, "Sensor selection in distributed multiple-radar architectures for localization: A knapsack problem formulation," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 247–260, Jan. 2012.

[23] F. Yousefian, A. Nedić, and U. V. Shanbhag, "On stochastic gradient and subgradient methods with adaptive steplength sequences," *Automatica*, vol. 48, no. 1, pp. 56–67, Jan. 2012.

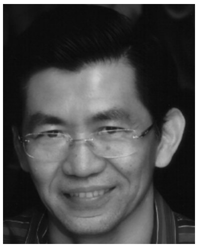
[24] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.

[25] S. Magnússon, P. C. Weeraddana, M. G. Rabbat, and C. Fischione, "On the convergence of alternating direction Lagrangian methods for nonconvex structured optimization problems," *IEEE Trans. Control Netw. Syst.*, vol. PP, no. 99, pp. 1–14, doi: 10.1109/TCNS.2015.2476198, preprint. Q2

[26] B. T. Polyak, *Introduction to Optimization*. New York, NY, USA: Optimization Software, Inc., 1987.

**Ying Ma** received the B.E. degree in electronic engineering from the Changchun Institute of Technology, Jin Lin, China, in 2011. She is currently working toward the Ph.D. degree with the Research Institute of Communication Technology, Beijing Institute of Technology, Beijing, China. Since July 2015, she has been with the Department of Electrical and Computer Engineering, University of Southampton, Southampton, U.K., where she is a visiting Ph.D. student under the supervision of Prof. L. Hanzo and Prof. S. Chen. Her research interests include distributed computation, multiple-input multiple-output systems, and cooperative communication.



**Sheng Chen** (M'90–SM'97–F'08) received the B.Eng. degree in control engineering from the East China Petroleum Institute, Dongying, China, in January 1982, the Ph.D. degree in control engineering from the City University, London, U.K., in September 1986, and the D.Sc. degree from the University of Southampton, Southampton, U.K., in 2005. He held research and academic appointments with the University of Sheffield, the University of Edinburgh, and the University of Portsmouth, U.K., from 1986 to 1999. Since 1999, he has been with the Electronics and Computer Science Department, University of Southampton, where he is a Professor of intelligent systems and signal processing. His research interests include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, evolutionary computation methods, and optimization. He has authored more than 550 research papers. He is a Fellow of the United Kingdom Royal Academy of Engineering, a Fellow of the IET, and a Distinguished Adjunct Professor with King Abdulaziz University, Jeddah, Saudi Arabia. He was an ISI Highly Cited Researcher in engineering in 2004.

**Chengwen Xing** (S'08–M'10) received the B.Eng. degree from Xidian University, Xi'an, China, in 2005, and the Ph.D. degree from the University of Hong Kong, Hong Kong, in 2010. Since September 2010, he has been with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China, where he is currently an Associate Professor. From September 2012 to December 2012, he was a Visiting Scholar at the University of Macau. His current research interests include statistical signal processing, convex optimization, multivariate statistics, combinatorial optimization, massive MIMO systems, and high-frequency band communication systems. He is currently serving as an Associate Editor for the IEEE Transactions on Vehicular Technology, *KSII Transactions on Internet and Information Systems*, *Transactions on Emerging Telecommunications Technologies*, and *China Communications*.

**Xiangyuan Bu** received the B.Eng., Master's, and Ph.D. degrees from the Beijing Institute of Technology, Beijing, China, in 1987, 1993, and 2007, respectively. He is currently a Professor in the Wireless Communications and Networks Laboratory, School of Information and Electronics, Beijing Institute of Technology. From July 1987 to September 1990, he was a Researcher in the Institute of Harbin, 674 Factory, and from April 1993 to May 2002, he was a Researcher in the No. 23 Institute of China Aerospace Science and Industry. His current research interests include wireless communication and signal processing.

**Lajos Hanzo** (F'04) received the D.Sc. degree in electronics in 1976 and the Doctorate degree in 1983. During his 38-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, U.K., where he holds the Chair in telecommunications. He has successfully supervised 100+ Ph.D. students, coauthored 20 John Wiley/IEEE Press books on mobile radio communications, totalling in excess of 10 000 pages, published 1590 research entries at IEEE Xplore, acted both as the TPC Chair and the General Chair of IEEE conferences, presented keynote lectures, and has received a number of distinctions. He is currently directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council, U.K., the European Research Council's Advanced Fellow Grant, and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. In 2009, he received the honorary doctorate "Doctor Honoris Causa" by the Technical University of Budapest. During 2008–2012, he was the Editor-in-Chief of the IEEE Press and also a Chaired Professor with Tsinghua University, Beijing, China. His research is funded by the European Research Council's Senior Research Fellow Grant. He is a Fellow of the Royal Academy of Engineering, IET, and EURASIP. For further information on research in progress and associated publications, please refer to http://www-mobile.ecs.soton.ac.uk. He has 25 000+ citations.

**QUERIES**

Q1. Author: “Scenario” is spelled as “senario.” Please check.

Q2. Author: Please provide full bibliographic details in Ref. [25].