Fast Fractional Programming for Multi-Cell Integrated Sensing and Communications

Yannan Chen†, Yi Feng†, Xiaoyang Li§†, Licheng Zhao§, and Kaiming Shen†
†School of Science and Engineering, The Chinese University of Hong Kong (Shenzhen), China
§Shenzhen Research Institute of Big Data, China
E-mail: {yannanchen, yifeng}@link.cuhk.edu.cn; {lixiaoyang, zhaolicheng}@sribd.cn; shenkaiming@cuhk.edu.cn

Abstract*—This paper concerns the coordinate multi-cell beamforming design for integrated sensing and communications (ISAC). The optimization objective is to maximize a weighted sum of the communication rates and the sensing Fisher information. The conventional beamforming method for massive multiple-input multiple-output transmission, i.e., the weighted minimum mean square error (WMMSE) algorithm, has a natural extension to the ISAC problem scenario from a fractional programming (FP) perspective. However, the extended WMMSE algorithm requires computing the $N \times N$ matrix inverse extensively, where N is proportional to the antenna array size, so the algorithm becomes quite costly when antennas are massively deployed. To address this issue, we develop a nonhomogeneous bound and use it in conjunction with the FP technique to solve the ISAC beamforming problem without the need to invert any large matrices. It is further shown that the resulting new FP algorithm has an intimate connection with gradient projection, based on which the convergence can be accelerated.

I. INTRODUCTION

Integrated sensing and communications (ISAC) is an emerging wireless technique that reuses the network infrastructure and radio signals for both communications and sensing—which used to be dealt with separately in the conventional networks, in order to reduce the infrastructure cost and boost the spectral efficiency. This work focuses on the large antenna array case of ISAC, aiming at a system-level optimization by coordinating the antenna beamformers across multiple cells.

The ISAC beamforming problem is a nontrivial task due to the nonconvex nature of the underlying optimization problem. Quite a few advanced optimization tools have been considered in the previous attempts. Semi-definite relaxation (SDR) is a typical example because the ISAC beamforming problem can be somehow relaxed in a quadratic form, e.g., as a quadratic semi-definite program (QSDP) [2] or a semi-definite program (SDP) [3]. Successive convex approximation (SCA) constitutes another popular approach in this area. For example, [4] uses SCA to convert the ISAC beamforming design to a second-order cone programming (SOCP) problem—which can be efficiently solved by the standard convex optimization method. Another line of studies [5], [6] utilize the majorization-minimization (MM) theory to make the ISAC beamforming problem convex, especially when the passive beamforming of the intelligent reflecting surface (IRS) is involved. Moreover, because the ISAC beamforming problem is fractionally structured, the fractional programming (FP) technique, i.e., the quadratic transform [7], [8], forms the building block of [9], [10].

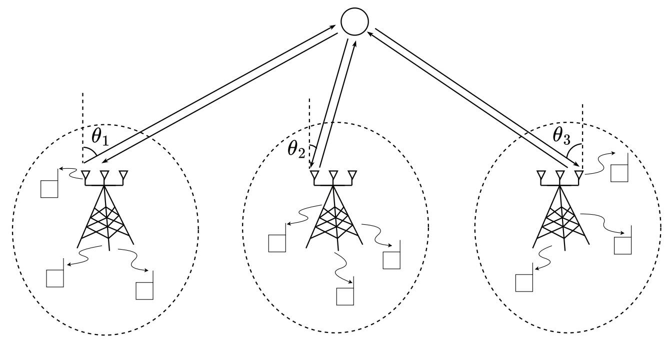


Fig. 1. A multi-cell ISAC system with $L = 3$ and $K = 3$. The circle is the point target to sense. The arrows are the transmit signals and the echo signals.

To exploit the degrees-of-freedom (DoF), the massive MIMO technology has been considered for the ISAC system. In principle, as shown in [11], the efficiency of target detection improves with the antenna array size. From the algorithm design viewpoint, however, the large antenna array can pose a tough challenge. Actually, even considering the communications task alone, the beamforming design with massive antennas is already quite difficult. For instance, although the WMMSE algorithm [12], [13] has been extensively used for the MIMO transmission, it is no longer suited for the massive MIMO case because the algorithm then entails computing the large matrix inverse extensively per iteration. Other standard optimization methods such as SDR and SCA are faced with similar issues since they involve the matrix inverse operation implicitly when performing the interior-point optimization. Some recent efforts aim to improve the efficiency of WMMSE in the massive MIMO case, e.g., [14] proposes a light WMMSE algorithm that has a lower complexity of the matrix inverse computation under certain conditions. The present paper is most closely related to a series of recent works [15], [16] that use a nonhomogeneous bound from [17] to avoid matrix inverse. While [15], [16] focus on the transmit beamforming alone, this work proposes a novel use of the

The complete version [1] is available at https://arxiv.org/abs/2406.10910. This work was supported in part by Guangdong Major Project of Basic and Applied Basic Research (No. 2023B0303000001), in part by Shenzhen Steady Funding Program, in part by National Nature Science Foundation of China (No. 62206182), and in part by Guangdong Basic and Applied Basic Research Foundation (No. 2024A1515010154).

non-homogeneous bound that accounts for both the transmit beamforming and the echo receive beamforming.

II. Multi-Cell ISAC System Model

Consider a downlink multi-cell ISAC system with L cells. Assume that each cell has K downlink users. For each BS $\ell = 1, 2, \ldots, L$, there are two tasks: (i) send independent messages to the K downlink users in its cell by spatial multiplexing; (ii) detect the direction of arrival (DoA), θ_l, for a point target. An illustrative example of the ISAC system as described above is shown in Fig. 1. We assume that BS ℓ has N^t_ℓ transmit antennas as well as N^r_ℓ receive antennas for detecting the echos. The kth downlink user in the ℓth cell is indexed as (ℓ, k). User (ℓ, k) has $M_{\ell k}$ receive antennas. Notably, the transmit antenna array size N^t_ℓ and the echo antenna array size N^r_ℓ at the BS side can be large, whereas the receive antenna array size $M_{\ell k}$ at the terminal side is often small, as typically assumed for a massive MIMO network.

A. Communications Model

Let $\boldsymbol{S}_{\ell k} \in \mathbb{C}^{d_{\ell k}\times T}$ be the normalized symbol sequence intended for user (ℓ, k), where $d_{\ell k}$ is the number of data streams and T is the block length. Note that $d_{\ell k}$ is often small since $d_{\ell k} \le M_{\ell k}$. Let $\boldsymbol{W}_{\ell k} \in \mathbb{C}^{N^t_\ell \times d_{\ell k}}$ be the transmit beamformer of BS ℓ for $\boldsymbol{S}_{\ell k}$. The different data streams are assumed to be statistically independent. Denote by $\boldsymbol{H}_{\ell k,i} \in \mathbb{C}^{M_{\ell k}\times N^t_i}$ the channel from BS i to user (ℓ, k). Each entry of the background noise $\boldsymbol{\Delta}_{\ell k} \in \mathbb{C}^{M_{\ell k}\times T}$ is drawn i.i.d. from $\mathcal{CN}(0, \sigma^2)$. The data rate for user (ℓ, k) can be computed as

$$R_{\ell k} = \log \left| \boldsymbol{I}_{d_{\ell k}} + \boldsymbol{W}^{\mathrm{H}}_{\ell k}\boldsymbol{H}^{\mathrm{H}}_{\ell k,\ell}\boldsymbol{F}^{-1}_{\ell k}\boldsymbol{H}_{\ell k,\ell}\boldsymbol{W}_{\ell k} \right|, \tag{1}$$

where

$$\boldsymbol{F}_{\ell k} = \sum_{(i,j)\neq(\ell,k)} \boldsymbol{H}_{\ell k,i}\boldsymbol{W}_{ij}\boldsymbol{W}^{\mathrm{H}}_{ij}\boldsymbol{H}^{\mathrm{H}}_{\ell k,i} + \sigma^2\boldsymbol{I}_{M_{\ell k}}. \tag{2}$$

For the good of communications, we wish to maximize the data rates throughout the network.

B. Sensing Model

Recall that BS ℓ has N^t_ℓ transmit antennas (for communications) and N^r_ℓ receive antennas (for sensing). Denote by $\boldsymbol{a}^t_\ell(\theta_\ell)$ the steering vector of the transmit antennas, $\boldsymbol{a}^r_\ell(\theta_\ell)$ the steering vector of the receive antennas, and $\xi_{\ell i}$ the reflection coefficient from BS i to BS ℓ. The echo signal is corrupted by the background noise $\widetilde{\boldsymbol{\Delta}}_\ell \in \mathbb{C}^{N^r_\ell\times T}$ drawn i.i.d. from $\mathcal{CN}(0, \tilde{\sigma}^2)$. Each BS ℓ recovers the DoA θ_ℓ from the received echo signal $\widetilde{\boldsymbol{Y}}_\ell$.

Mean squared error (MSE) is a common performance metric of estimation. Nevertheless, it is difficult to analyze the MSE of θ_ℓ in our problem case. Instead, we adopt the Fisher information as the performance metric, following the previous works [18], [19] in the ISAC field. Define the response matrix

$$\boldsymbol{G}_{\ell\ell} = \xi_{\ell\ell}\boldsymbol{a}^r_\ell(\theta_\ell)(\boldsymbol{a}^t_\ell(\theta_\ell))^\top. \tag{3}$$

Denote by $\boldsymbol{G}_{\ell i}$ the interference channel from BS i to BS ℓ for any $i \neq \ell$. Let

$$\boldsymbol{\mu}_\ell = \sum_{i=1,i\neq\ell}^{L}\sum_{j=1}^{K}(\boldsymbol{I}_T \otimes \boldsymbol{G}_{\ell i}\boldsymbol{W}_{ij})\boldsymbol{s}_{ij} + \tilde{\boldsymbol{\delta}}_\ell, \tag{4}$$

where $\tilde{\boldsymbol{\delta}}_\ell = \mathrm{vec}(\widetilde{\boldsymbol{\Delta}}_\ell)$ and $\boldsymbol{s}_{\ell k} = \mathrm{vec}(\boldsymbol{S}_{\ell k})$. Define $\boldsymbol{Q}_\ell$ to be

$$\begin{aligned}\boldsymbol{Q}_\ell &= \mathbb{E}\left[\boldsymbol{\mu}_\ell\boldsymbol{\mu}^{\mathrm{H}}_\ell\right] \\ &= \boldsymbol{I}_T \otimes \left(\sum_{i=1,i\neq\ell}^{L}\sum_{j=1}^{K}\boldsymbol{G}_{\ell i}\boldsymbol{W}_{ij}\boldsymbol{W}^{\mathrm{H}}_{ij}\boldsymbol{G}^{\mathrm{H}}_{\ell i} + \tilde{\sigma}^2\boldsymbol{I}_{N^r_\ell}\right).\end{aligned}$$

The Fisher information of θ_ℓ is then given by [20]

$$J_\ell = 2T\left\{\sum_{k=1}^{K}\mathrm{tr}\left((\dot{\boldsymbol{G}}_{\ell\ell}\boldsymbol{W}_{\ell k})^{\mathrm{H}}\widehat{\boldsymbol{Q}}^{-1}_\ell(\dot{\boldsymbol{G}}_{\ell\ell}\boldsymbol{W}_{\ell k})\right)\right\}, \tag{5}$$

where $\dot{\boldsymbol{G}}_{\ell\ell} = \partial\boldsymbol{G}_{\ell\ell}/\partial\theta_\ell$ and

$$\widehat{\boldsymbol{Q}}_\ell = \sum_{i\neq\ell}\sum_{j=1}^{K}\boldsymbol{G}_{\ell i}\boldsymbol{W}_{ij}\boldsymbol{W}^{\mathrm{H}}_{ij}\boldsymbol{G}^{\mathrm{H}}_{\ell i} + \tilde{\sigma}^2\boldsymbol{I}_{N^r_\ell}. \tag{6}$$

For the good of sensing, we wish to maximize the Fisher information of θ_ℓ at each BS ℓ.

C. ISAC Beamforming Problem

To account for both communications and sensing, we consider maximizing a weighted sum of data rates and Fisher information:

$$\underset{\underline{\boldsymbol{W}}}{\mathrm{maxmize}} \quad \sum_{\ell=1}^{L}\sum_{k=1}^{K}\omega_{\ell k}R_{\ell k} + \sum_{\ell=1}^{L}\beta_\ell J_\ell \tag{7a}$$

$$\text{subject to} \quad \sum_{k=1}^{K}\|\boldsymbol{W}_{\ell k}\|^2_F \le P, \tag{7b}$$

where $\underline{\boldsymbol{W}} = \{\boldsymbol{W}_{11}, \ldots, \boldsymbol{W}_{LK}\}$ is the collocation of all beamformers, $\omega_{\ell k}, \beta_\ell \ge 0$ are the given nonnegative weights reflecting the priorities of the communications and sensing tasks, and P is the power budget of each BS.

III. Main Results

As a key observation, problem (7) is fractionally structured. In light of this FP interpretation, we first show that the traditional WMMSE algorithm [12], [13] for the communication beamforming can be extended for the ISAC beamforming.

A. Extended WMMSE for ISAC Beamforming

First, we use the Lagrangian dual transform [8] to move the ratios to the outside of log-determinants for the term $\sum_{\ell=1}^{L}\sum_{k=1}^{K}\omega_{\ell k}R_{\ell k}$ in (7a). Problem (7) is converted to

$$\underset{\underline{\boldsymbol{W}},\underline{\boldsymbol{\Gamma}}}{\mathrm{maximize}} \quad f_r(\underline{\boldsymbol{W}}, \underline{\boldsymbol{\Gamma}}) + \sum_{\ell=1}^{L}\beta_\ell J_\ell \tag{8a}$$

$$\text{subject to} \quad \sum_{k=1}^{K}\|\boldsymbol{W}_{\ell k}\|^2_F \le P \tag{8b}$$

$$f_q(\underline{\boldsymbol{W}},\underline{\boldsymbol{\Gamma}},\underline{\boldsymbol{Y}},\underline{\widetilde{\boldsymbol{Y}}})=\sum_{\ell,k}\Big[\mathrm{tr}\big(2\Re\{\boldsymbol{W}_{\ell k}^{\mathrm{H}}\boldsymbol{\Lambda}_{\ell k}\}-\omega_{\ell k}\boldsymbol{Y}_{\ell k}^{\mathrm{H}}\boldsymbol{U}_{\ell k}\boldsymbol{Y}_{\ell k}(\boldsymbol{I}+\boldsymbol{\Gamma}_{\ell k})-2T\beta_\ell\widetilde{\boldsymbol{Y}}_{\ell k}^{\mathrm{H}}\widehat{\boldsymbol{Q}}_\ell\widetilde{\boldsymbol{Y}}_{\ell k}\big)+\omega_{\ell k}\big(\log|\boldsymbol{I}+\boldsymbol{\Gamma}_{\ell k}|-\mathrm{tr}(\boldsymbol{\Gamma}_{\ell k})\big)\Big] \tag{11}$$

where

$$f_r(\underline{\boldsymbol{W}},\underline{\boldsymbol{\Gamma}})=\sum_{\ell=1}^{L}\sum_{k=1}^{K}\omega_{\ell k}\Big[\log\big|\boldsymbol{I}_{d_{\ell k}}+\boldsymbol{\Gamma}_{\ell k}\big|-\mathrm{tr}\big(\boldsymbol{\Gamma}_{\ell k}\big)+\mathrm{tr}\Big((\boldsymbol{I}_{d_{\ell k}}+\boldsymbol{\Gamma}_{\ell k})\boldsymbol{W}_{\ell k}^{\mathrm{H}}\boldsymbol{H}_{\ell k,\ell}^{\mathrm{H}}\boldsymbol{U}_{\ell k}^{-1}\boldsymbol{H}_{\ell k,\ell}\boldsymbol{W}_{\ell k}\Big)\Big] \tag{9}$$

with $\boldsymbol{U}_{\ell k}=\sum_{i=1}^{L}\sum_{j=1}^{K}\boldsymbol{H}_{\ell k,i}\boldsymbol{W}_{ij}\boldsymbol{W}_{ij}^{\mathrm{H}}\boldsymbol{H}_{\ell k,i}^{\mathrm{H}}+\sigma^2\boldsymbol{I}_{M_{\ell k}}$. For fixed $\underline{\boldsymbol{W}}$, problem (8) is convex in $\underline{\boldsymbol{\Gamma}}$. According to the first-order condition, each $\boldsymbol{\Gamma}_{\ell k}$ can be optimally determined as

$$\boldsymbol{\Gamma}_{\ell k}^{\star}=\boldsymbol{W}_{\ell k}^{\mathrm{H}}\boldsymbol{H}_{\ell k,\ell}^{\mathrm{H}}\boldsymbol{F}_{\ell k}^{-1}\boldsymbol{H}_{\ell k,\ell}\boldsymbol{W}_{\ell k}. \tag{10}$$

Notice that (8) is a sum-of-weighted-ratios problem of $\underline{\boldsymbol{W}}$ when $\underline{\boldsymbol{\Gamma}}$ is fixed, so the quadratic transform [8] is applicable. The optimization objective in (8a) is then further recast to $f_q(\underline{\boldsymbol{W}},\underline{\boldsymbol{\Gamma}},\underline{\boldsymbol{Y}},\underline{\widetilde{\boldsymbol{Y}}})$ as shown in (11) with

$$\boldsymbol{\Lambda}_{\ell k}=\omega_{\ell k}\boldsymbol{H}_{\ell k,\ell}^{\mathrm{H}}\boldsymbol{Y}_{\ell k}(\boldsymbol{I}_{d_{\ell k}}+\boldsymbol{\Gamma}_{\ell k})+2T\beta_\ell\dot{\boldsymbol{G}}_{\ell\ell}^{\mathrm{H}}\widetilde{\boldsymbol{Y}}_{\ell k}. \tag{12}$$

For the notational clarity in (11), we use $\boldsymbol{Y}_{\ell k}$ ($\widetilde{\boldsymbol{Y}}_{\ell k}$) to denote each auxiliary variable introduced for the communication task (sensing task). The resulting reformulation of problem (8) is

$$\underset{\underline{\boldsymbol{W}},\underline{\boldsymbol{\Gamma}},\underline{\boldsymbol{Y}},\underline{\widetilde{\boldsymbol{Y}}}}{\text{maximize}}\quad f_q(\underline{\boldsymbol{W}},\underline{\boldsymbol{\Gamma}},\underline{\boldsymbol{Y}},\underline{\widetilde{\boldsymbol{Y}}}) \tag{13a}$$

$$\text{subject to}\quad \sum_{k=1}^{K}\|\boldsymbol{W}_{\ell k}\|_F^2\le P \tag{13b}$$

When $\underline{\boldsymbol{W}}$ and $\underline{\boldsymbol{\Gamma}}$ are both held fixed, the above problem is jointly convex in $\underline{\boldsymbol{Y}}$ and $\underline{\widetilde{\boldsymbol{Y}}}$, so we can optimally determine them by the first-order condition as

$$\boldsymbol{Y}_{\ell k}^{\star}=\boldsymbol{U}_{\ell k}^{-1}\boldsymbol{H}_{\ell k,\ell}\boldsymbol{W}_{\ell k}, \tag{14}$$

$$\widetilde{\boldsymbol{Y}}_{\ell k}^{\star}=\widehat{\boldsymbol{Q}}_{\ell}^{-1}\dot{\boldsymbol{G}}_{\ell\ell}\boldsymbol{W}_{\ell k}. \tag{15}$$

By the identity $\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})=\|\boldsymbol{A}\|_F^2$, we get the optimal $\boldsymbol{W}_{\ell k}$:

$$\boldsymbol{W}_{\ell k}^{\star}=\arg\min_{\underline{\boldsymbol{W}}\in\mathcal{W}}\|\boldsymbol{L}_\ell^{\frac{1}{2}}(\boldsymbol{W}_{\ell k}-\boldsymbol{L}_\ell^{-1}\boldsymbol{\Lambda}_{\ell k})\|_F^2, \tag{16}$$

where

$$\mathcal{W}=\left\{\underline{\boldsymbol{W}}:\sum_{k=1}^{K}\|\boldsymbol{W}_{\ell k}\|_F^2\le P,\ \forall\ell\right\}, \tag{17}$$

and

$$\boldsymbol{L}_\ell=\sum_{i=1}^{L}\sum_{j=1}^{K}\omega_{ij}\boldsymbol{H}_{ij,\ell}^{\mathrm{H}}\boldsymbol{Y}_{ij}(\boldsymbol{I}_{d_{ij}}+\boldsymbol{\Gamma}_{ij})\boldsymbol{Y}_{ij}^{\mathrm{H}}\boldsymbol{H}_{ij,\ell}+2T\sum_{i=1,i\neq\ell}^{L}\sum_{j=1}^{K}\boldsymbol{G}_{i\ell}^{\mathrm{H}}(\beta_i\widetilde{\boldsymbol{Y}}_{ij}\widetilde{\boldsymbol{Y}}_{ij}^{\mathrm{H}})\boldsymbol{G}_{i\ell}. \tag{18}$$

We summarize the above iterative optimization steps in Algorithm 1.

Algorithm 1 Extended WMMSE for ISAC Beamforming

1: Initialize $\underline{\boldsymbol{W}}$ to a feasible value.
2: **repeat**
3: Update each $\boldsymbol{\Gamma}_{\ell k}$ by (10).
4: Update each $\boldsymbol{Y}_{\ell k}$ and $\widetilde{\boldsymbol{Y}}_{\ell k}$ by (14) and (15), resp.
5: Update each $\boldsymbol{W}_{\ell k}$ by (16).
6: **until** the objective value converges

Remark 1: Notice that updating $\underline{\boldsymbol{W}}$ as in (16) and updating $\underline{\widetilde{\boldsymbol{Y}}}$ as in (15) both entail inverting matrices, which can be quite costly when massive antennas are deployed so that N_ℓ^r or/and N_ℓ^t is large.

B. Large Matrix Inverse Elimination

Before tackling the large matrix inverse issue of the ISAC beamforming problem, we first illustrate how this issue can be addressed in a toy example. The main tool is stated in the following lemma:

Lemma 1 (Nonhomogeneous Bound): Suppose that the two Hermitian matrices $\boldsymbol{L},\boldsymbol{K}\in\mathbb{C}^{d\times d}$ satisfy $\boldsymbol{L}\preceq\boldsymbol{K}$, e.g., when $\boldsymbol{K}=\lambda\boldsymbol{I}$ where $\lambda=\lambda_{\max}(\boldsymbol{L})$ with $\lambda_{\max}(\boldsymbol{L})$ being the largest eigenvalue of $\boldsymbol{L}$. Then for any two matrices $\boldsymbol{X},\boldsymbol{Z}\in\mathbb{C}^{d\times m}$, one has

$$\mathrm{tr}(\boldsymbol{X}^{\mathrm{H}}\boldsymbol{L}\boldsymbol{X})\le\mathrm{tr}\big(\boldsymbol{X}^{\mathrm{H}}\boldsymbol{K}\boldsymbol{X}+2\Re\{\boldsymbol{X}^{\mathrm{H}}(\boldsymbol{L}-\boldsymbol{K})\boldsymbol{Z}\}+\boldsymbol{Z}^{\mathrm{H}}(\boldsymbol{K}-\boldsymbol{L})\boldsymbol{Z}\big), \tag{19}$$

where the equality holds if $\boldsymbol{Z}=\boldsymbol{X}$.

Example 1: Consider the following single-ratio problem:

$$\underset{\boldsymbol{X}\in\mathcal{X}}{\text{maximize}}\quad \mathrm{tr}((\boldsymbol{A}\boldsymbol{X})^{\mathrm{H}}(\boldsymbol{B}\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}})^{-1}(\boldsymbol{A}\boldsymbol{X})), \tag{20}$$

where $\boldsymbol{A},\boldsymbol{B}\in\mathbb{C}^{n\times d}$, $\boldsymbol{X}\in\mathbb{C}^{d\times m}$, and $\mathcal{X}$ is a nonempty set. By quadratic transform, the above problem can be recast to

$$\underset{\boldsymbol{X}\in\mathcal{X},\boldsymbol{Y}}{\text{maximize}}\quad f_q(\boldsymbol{X},\boldsymbol{Y}), \tag{21}$$

where

$$f_q(\boldsymbol{X},\boldsymbol{Y})=\mathrm{tr}\big(2\Re\{\boldsymbol{X}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{Y}\}-\boldsymbol{X}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{X}\big). \tag{22}$$

Next, treating

$$\boldsymbol{L}=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{B} \tag{23}$$

and applying the nonhomogeneous bound in Lemma 1, we can convert problem (21) to

$$\underset{\boldsymbol{X}\in\mathcal{X},\boldsymbol{Y},\boldsymbol{Z}}{\text{maximize}}\quad g_o(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}), \tag{24}$$

$$g_s(\underline{\boldsymbol{W}},\underline{\boldsymbol{\Gamma}},\underline{\boldsymbol{Y}},\underline{\widetilde{\boldsymbol{Y}}},\underline{\boldsymbol{Z}},\underline{\widetilde{\boldsymbol{Z}}})=\sum_{\ell,k}\Big[\mathrm{tr}\big(\Re\{2\boldsymbol{W}_{\ell k}^{\mathrm{H}}\boldsymbol{\Lambda}_{\ell k}+2\widetilde{\boldsymbol{Y}}_{\ell k}^{\mathrm{H}}(\tilde{\lambda}_\ell\boldsymbol{I}_{N_\ell^r}-\widetilde{\boldsymbol{L}}_\ell)\widetilde{\boldsymbol{Z}}_{\ell k}+\widetilde{\boldsymbol{Z}}_{\ell k}^{\mathrm{H}}(\widetilde{\boldsymbol{L}}_\ell-\tilde{\lambda}_\ell\boldsymbol{I}_{N_\ell^r})\widetilde{\boldsymbol{Z}}_{\ell k}-(2\boldsymbol{W}_{\ell k}-\boldsymbol{Z}_{\ell k})^{\mathrm{H}}\boldsymbol{D}_\ell\boldsymbol{Z}_{\ell k}$$
$$+\lambda_\ell(2\boldsymbol{W}_{\ell k}^{\mathrm{H}}\boldsymbol{Z}_{\ell k}-\boldsymbol{Z}_{\ell k}^{\mathrm{H}}\boldsymbol{Z}_{\ell k}-\boldsymbol{W}_{\ell k}^{\mathrm{H}}\boldsymbol{W}_{\ell k})\}\big)+\omega_{\ell k}\log|\boldsymbol{I}_{d_{\ell k}}+\boldsymbol{\Gamma}_{\ell k}|-\mathrm{tr}\big(\omega_{\ell k}\boldsymbol{\Gamma}_{\ell k}+\omega_{\ell k}\sigma^2(\boldsymbol{I}_{d_{\ell k}}+\boldsymbol{\Gamma}_{\ell k})\boldsymbol{Y}_{\ell k}^{\mathrm{H}}\boldsymbol{Y}_{\ell k}\big)\Big] \quad (34)$$

where the new objective function is

$$g_o(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z})=\mathrm{tr}\Big(2\Re\{\boldsymbol{X}^{\mathrm{H}}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{Y})+\boldsymbol{X}^{\mathrm{H}}(\lambda\boldsymbol{I}_d-\boldsymbol{L})\boldsymbol{Z}\}+\boldsymbol{Z}^{\mathrm{H}}(\boldsymbol{L}-\lambda\boldsymbol{I}_d)\boldsymbol{Z}-\lambda\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X}\Big) \quad (25)$$

with $\lambda=\lambda_{\max}(\boldsymbol{L})$. $\boldsymbol{X}$ can now be optimally determined for $g_o(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z})$ in closed form without computing the matrix inverse when $(\boldsymbol{Y},\boldsymbol{Z})$ are held fixed. This desirable result motivates us to apply Lemma 1 one more time to get rid of the matrix inverse for the optimal update of $\boldsymbol{Y}$ in solving (24). Specifically, rewriting (25) as

$$g_o(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}^\star)=\mathrm{tr}\Big(\Re\{-\boldsymbol{Y}^{\mathrm{H}}(\boldsymbol{B}\boldsymbol{Z}^\star(2\boldsymbol{X}-\boldsymbol{Z}^\star)^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}})\boldsymbol{Y}+2\boldsymbol{Y}^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{X})+\lambda(2\boldsymbol{X}^{\mathrm{H}}\boldsymbol{Z}^\star-\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X}-(\boldsymbol{Z}^\star)^{\mathrm{H}}\boldsymbol{Z}^\star)\}\Big) \quad (26)$$

and treating

$$\widetilde{\boldsymbol{L}}=\boldsymbol{B}\boldsymbol{Z}^\star(2\boldsymbol{X}-\boldsymbol{Z}^\star)^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}, \quad (27)$$

problem (24) can be further recast to

$$\underset{\boldsymbol{X}\in\mathcal{X},\boldsymbol{Y},\boldsymbol{Z}^\star,\widetilde{\boldsymbol{Z}}}{\text{maximize}}\quad g_s(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}^\star,\widetilde{\boldsymbol{Z}}), \quad (28)$$

where

$$g_s(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}^\star,\widetilde{\boldsymbol{Z}})=\mathrm{tr}\big(\Re\{2\boldsymbol{Y}^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{X}+(\tilde{\lambda}\boldsymbol{I}_n-\widetilde{\boldsymbol{L}})\widetilde{\boldsymbol{Z}})+\widetilde{\boldsymbol{Z}}^{\mathrm{H}}(\widetilde{\boldsymbol{L}}-\tilde{\lambda}\boldsymbol{I}_n)\widetilde{\boldsymbol{Z}}+\lambda(2\boldsymbol{X}^{\mathrm{H}}\boldsymbol{Z}^\star-\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X}-(\boldsymbol{Z}^\star)^{\mathrm{H}}\boldsymbol{Z}^\star)-\tilde{\lambda}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{H}}\}\big) \quad (29)$$

with $\tilde{\lambda}=\lambda_{\max}(\widetilde{\boldsymbol{L}})$. We propose optimizing the variables of $g_s(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z},\widetilde{\boldsymbol{Z}})$ in an iterative fashion as

$$\cdots\to\boldsymbol{X}^\tau\to\boldsymbol{Z}^\tau\to\boldsymbol{Y}^\tau\to\widetilde{\boldsymbol{Z}}^\tau\to\boldsymbol{X}^{\tau+1}\to\cdots.$$

We now specify the iterative optimization steps. First, according to Lemma 1, $\boldsymbol{Z}$ and $\widetilde{\boldsymbol{Z}}$ are optimally updated as

$$\boldsymbol{Z}^\star=\boldsymbol{X}\quad\text{and}\quad\widetilde{\boldsymbol{Z}}^\star=\boldsymbol{Y}. \quad (30)$$

With the optimal $\widetilde{\boldsymbol{Z}}^\star=\boldsymbol{Y}$ plugged in $g_s(\boldsymbol{X},\boldsymbol{Y},\widetilde{\boldsymbol{Z}},\boldsymbol{Z})$, we can find the optimal update of $\boldsymbol{X}$ by completing the square as

$$\boldsymbol{X}^\star=\boldsymbol{Z}+\frac{1}{\lambda}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{Y}-\boldsymbol{L}\boldsymbol{Z}). \quad (31)$$

Likewise, after $\boldsymbol{Z}$ has been optimally updated to $\boldsymbol{X}$, we can find the optimal update of $\boldsymbol{Y}$ by completing the square as

$$\boldsymbol{Y}^\star=\widetilde{\boldsymbol{Z}}+\frac{1}{\tilde{\lambda}}(\boldsymbol{A}\boldsymbol{X}-\widetilde{\boldsymbol{L}}\widetilde{\boldsymbol{Z}}). \quad (32)$$

We remark that the above iterative optimization steps do not incur any matrix inverse.

C. Nonhomogeneous FP for ISAC Beamforming

We now extend the result of Example 1 to the multi-ratio FP case for the ISAC beamforming. First, we still apply the Lagrangian dual transform and the quadratic transform to problem (7), so that the original problem is converted to (13).

In order to get rid of large matrix inverse, we follow the procedure in Example 1 and reformulate problem (13) further. After two uses of the nonhomogeneous bound, problem (13) is equivalent to

$$\text{maximize}\quad g_s(\underline{\boldsymbol{W}},\underline{\boldsymbol{\Gamma}},\underline{\boldsymbol{Y}},\underline{\widetilde{\boldsymbol{Y}}},\underline{\boldsymbol{Z}},\underline{\widetilde{\boldsymbol{Z}}}) \quad (33a)$$
$$\text{subject to}\quad \sum_{k=1}^{K}\|\boldsymbol{W}_{\ell k}\|_F^2\le P, \quad (33b)$$

The new objective function is shown in (34) as displayed at the top of the page, where

$$\boldsymbol{D}_\ell=\sum_{i=1}^{L}\sum_{j=1}^{K}\omega_{\ell j}\boldsymbol{H}_{ij,\ell}^{\mathrm{H}}\boldsymbol{Y}_{ij}(\boldsymbol{I}_{d_{ij}}+\boldsymbol{\Gamma}_{ij})\boldsymbol{Y}_{ij}^{\mathrm{H}}\boldsymbol{H}_{ij,\ell}, \quad (35)$$

$$\widetilde{\boldsymbol{L}}_\ell=2T\sum_{i=1,i\neq\ell}^{L}\sum_{j=1}^{K}\boldsymbol{G}_{i\ell}(\beta_i\boldsymbol{Z}_{ij}(2\boldsymbol{W}_{ij}-\boldsymbol{Z}_{ij})^{\mathrm{H}})\boldsymbol{G}_{i\ell}^{\mathrm{H}}+\tilde{\sigma}^2\boldsymbol{I}_{N_\ell^r}, \quad (36)$$

$\tilde{\lambda}_\ell=\lambda_{\max}(\widetilde{\boldsymbol{L}}_\ell)$, $\lambda_\ell=\lambda_{\max}(\boldsymbol{L}_\ell)$, and $\boldsymbol{L}_\ell$ was shown earlier in (18). Now, following the steps in Example 1, we consider optimizing the variables in (33) iteratively as

$$\cdots\to\underline{\boldsymbol{W}}\to\underline{\boldsymbol{Z}}\to\underline{\boldsymbol{\Gamma}}\to\underline{\boldsymbol{Y}}\to\underline{\widetilde{\boldsymbol{Y}}}\to\underline{\widetilde{\boldsymbol{Z}}}\to\underline{\boldsymbol{W}}\to\cdots.$$

According to Lemma 1, $\underline{\boldsymbol{Z}}$ and $\underline{\widetilde{\boldsymbol{Z}}}$ are optimally updated as

$$\boldsymbol{Z}_{\ell k}^\star=\boldsymbol{W}_{\ell k}, \quad (37)$$
$$\widetilde{\boldsymbol{Z}}_{\ell k}^\star=\widetilde{\boldsymbol{Y}}_{\ell k}. \quad (38)$$

The optimal $\underline{\boldsymbol{\Gamma}}$ and $\underline{\boldsymbol{Y}}$ are still determined as in (10) and (14); notice that updating $\boldsymbol{Y}_{\ell k}$ requires computing the matrix inverse $\boldsymbol{U}_{\ell k}\in\mathbb{C}^{M_{\ell k}\times M_{\ell k}}$, but this is tolerable since the number of receiver antennas $M_{\ell k}$ is typically a small integer. With the optimal $\underline{\widetilde{\boldsymbol{Z}}}=\underline{\widetilde{\boldsymbol{Y}}}$ plugged in $g_s(\underline{\boldsymbol{W}},\underline{\boldsymbol{\Gamma}},\underline{\boldsymbol{Y}},\underline{\widetilde{\boldsymbol{Y}}},\underline{\widetilde{\boldsymbol{Z}}},\underline{\boldsymbol{Z}})$, we can find the optimal update of $\boldsymbol{W}_{\ell k}$ as

$$\boldsymbol{W}_{\ell k}^\star=\begin{cases}\widehat{\boldsymbol{W}}_{\ell k} & \text{if } \sum_j\|\widehat{\boldsymbol{W}}_{\ell j}\|_F^2\le P\\ \sqrt{\frac{P_\ell}{\sum_j\|\widehat{\boldsymbol{W}}_{\ell j}\|_F^2}}\widehat{\boldsymbol{W}}_{\ell k} & \text{otherwise,}\end{cases} \quad (39)$$

where

$$\widehat{\boldsymbol{W}}_{\ell k}=\boldsymbol{Z}_{\ell k}+\frac{1}{\lambda_\ell}(\boldsymbol{\Lambda}_{\ell k}-\boldsymbol{L}_\ell\boldsymbol{Z}_{\ell k}). \quad (40)$$

After $\underline{\boldsymbol{Z}}$ has been optimally updated to $\underline{\boldsymbol{W}}$, we obtain the

Algorithm 2 Nonhomogeneous FP for ISAC Beamforming

Initialize $\underline{\boldsymbol{W}}$ to a feasible value, $\underline{\widetilde{\boldsymbol{Y}}}$ according to (15), and $\underline{\widetilde{\boldsymbol{Z}}}$ according to (37).
repeat
  Update each $\boldsymbol{Z}_{\ell k}$ by (38).
  Update each $\boldsymbol{\Gamma}_{\ell k}$ by (10).
  Update each $\boldsymbol{Y}_{\ell k}$ and $\widetilde{\boldsymbol{Y}}_{\ell k}$ by (14) and (41), resp.
  Update each $\widetilde{\boldsymbol{Z}}_{\ell k}$ by (37).
  Update each $\boldsymbol{W}_{\ell k}$ by (39).
until the objective value converges

Algorithm 3 Fast FP for ISAC Beamforming

Initialize $\underline{\boldsymbol{W}}$ to a feasible value, $\underline{\widetilde{\boldsymbol{Y}}}$ according to (15), and $\underline{\widetilde{\boldsymbol{Z}}}$ according to (37).
repeat
  Update $\underline{\boldsymbol{V}}$ according to (43) and set $\boldsymbol{W}_{\ell k} = \boldsymbol{V}_{\ell k}$.
  Update each $\boldsymbol{Z}_{\ell k}$ by (38).
  Update each $\boldsymbol{\Gamma}_{\ell k}$ by (10).
  Update each $\boldsymbol{Y}_{\ell k}$ and $\widetilde{\boldsymbol{Y}}_{\ell k}$ by (14) and (41), resp.
  Update each $\widetilde{\boldsymbol{Z}}_{\ell k}$ by (37).
  Update each $\boldsymbol{W}_{\ell k}$ by (39).
until the objective value converges

optimal update of $\underline{\widetilde{\boldsymbol{Y}}}$ as

$$\widetilde{\boldsymbol{Y}}^{\star}_{\ell k} = \widetilde{\boldsymbol{Z}}_{\ell k} + \frac{1}{\lambda_\ell}(2T\beta_\ell \dot{\boldsymbol{G}}_{\ell\ell}\boldsymbol{W}_{\ell k} - \widehat{\boldsymbol{Q}}_\ell \widetilde{\boldsymbol{Z}}_{\ell k}). \tag{41}$$

Algorithm 2 summarizes the above steps and is referred to as the nonhomogeneous FP method for the ISAC beamforming. Differing from the extended WMMSE algorithm in Algorithm 1, the above new algorithm does not require computing any large matrix inverse.

D. Proposed Fast FP for ISAC Beamforming

In this work, we not only aim to eliminate large matrix inverse in order to reduce the per-iteration complexity, but also seek to accelerate the convergence in iterations. Our approach is based upon the following crucial observation: Algorithm 2 has a deep connection with gradient projection.

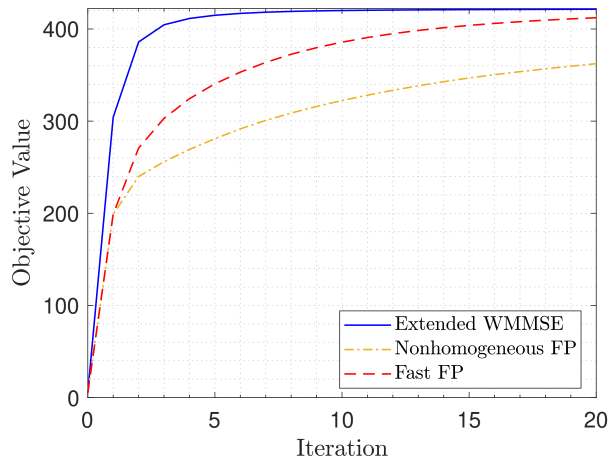
Proposition 1: Algorithm 2 is equivalent to a gradient projection:

$$\boldsymbol{W}^{\tau+1}_{\ell k} = \mathcal{P}_{\mathcal{W}}\left(\boldsymbol{W}^{\tau}_{\ell k} + \zeta^\tau \cdot \frac{\partial f_o(\underline{\boldsymbol{W}}^\tau)}{\partial \boldsymbol{W}_{\ell k}}\right), \tag{42}$$

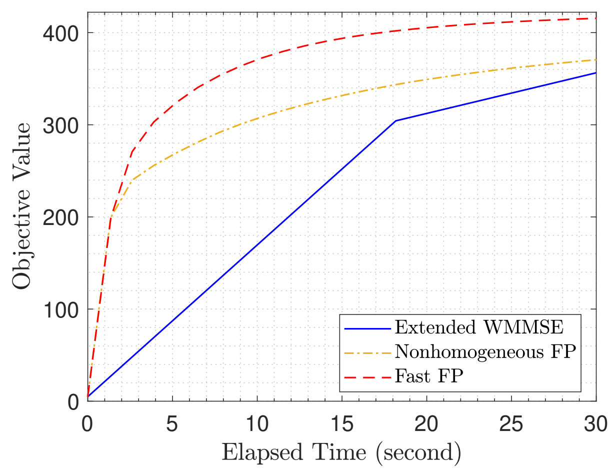
where τ is the iteration index, $\zeta^\tau > 0$ is the gradient step size in the τth iteration, $\mathcal{P}_{\mathcal{W}}(\cdot)$ is the Euclidean projection on $\mathcal{W}$, and $f_o(\underline{\boldsymbol{W}}^\tau)$ is the primal objective function in (7a).

Proof: Please see the long version of this paper [1]. ■

In light of the connection between Algorithm 2 and gradient projection, we can readily accelerate Algorithm 2 by using Nesterov's extrapolation strategy [21]. We now extrapolate each $\boldsymbol{W}_{\ell k}$ along the direction of the difference between the



(a) Objective value v.s. iteration number



(b) Objective value v.s. elapsed time

Fig. 2. The convergence behaviors of the different ISAC beamforming algorithms when $N_r = N_t = 128$.

preceding two iterates before the gradient projection, i.e.,

$$\boldsymbol{V}^{\tau-1}_{\ell k} = \boldsymbol{W}^{\tau-1}_{\ell k} + \upsilon^{\tau-1}(\boldsymbol{W}^{\tau-1}_{\ell k} - \boldsymbol{W}^{\tau-2}_{\ell k}), \tag{43}$$

$$\boldsymbol{W}^{\tau}_{\ell k} = \mathcal{P}_{\mathcal{W}}\left(\boldsymbol{V}^{\tau-1}_{\ell k} + \frac{1}{\lambda^\tau_\ell}(\boldsymbol{\Lambda}_{\ell k} - \boldsymbol{L}_\ell \boldsymbol{V}^{\tau-1}_{\ell k})\right), \tag{44}$$

where τ is the iteration index, the extrapolation step υ_τ is chosen as

$$\upsilon^\tau = \max\left\{\frac{\tau-2}{\tau+1}, 0\right\}, \quad \text{for } \tau = 1, 2, \ldots, \tag{45}$$

the starting point is $\underline{\boldsymbol{W}}^{-1} = \underline{\boldsymbol{W}}^{0}$. Algorithm 3 summarizes the above steps and is referred to as the fast FP algorithm.

IV. Numerical Results

We validate the performance of the proposed algorithms numerically in a 7-cell wrapped-hexagonal-around network. Within each cell, the BS is located at the center and 45 downlink users are randomly distributed. The BS-to-BS distance is set to be 800 meters. Let every BS have 128 transmit (resp. receive) antennas. Let every downlink user have 4 antennas; 4 data streams are intended for each of them. The maximum transmit power P at each BS is set to 20 dBm, the background noise power level at the downlink user side σ^2 is -80 dBm, and the background noise power level at the BS side $\tilde{\sigma}^2$ is set to 0 dBm. The downlink distance-dependent path-loss is given by $15.3 + 37.6\log_{10}(d) + \xi$ (in dB), where d represents the BS-to-user distance in meters, and ξ is a zero-mean Gaussian random variable with a standard variance of 8 dB—which models the shadowing effect. For the sensing, every reflection coefficient $\xi_{\ell i}$ is set to 1 as in [22]. The block length T equals 30. The priority weights in problem (7) are set as $\beta_\ell = 10^{-11}$ and $\omega_{\ell k} = 1$. We use the same starting point for all the competitor algorithms for the comparison fairness.

Fig. 2 shows the convergence behaviors of the different algorithms. Observe from Fig. 2(a) that the extended WMMSE converges faster than the nonhomogeneous FP and the fast FP in iterations. From an MM perspective, this is because the extended WMMSE gives a tighter approximation of the original objective function $f_o(\underline{\boldsymbol{W}})$. However, this does not imply that the absolute running time of the extended WMMSE is the shortest, because in the meanwhile the extended WMMSE requires heavier computations per iteration due to the large matrix inversion. Actually, it can be seen from Fig. 2(b) that the nonhomogeneous FP converges faster than extended WMMSE in time. For example, extended WMMSE requires about 23 seconds to reach the objective value of 158, while the nonhomogeneous FP merely requires 18 seconds. Thus, in practice, the nonhomogeneous FP still runs much faster than the extended WMMSE. Further, observe also that the fast FP outperforms the nonhomogeneous FP not only in terms of the iteration efficiency but also in terms of the time efficiency. For example, to reach the objective value of 160, the nonhomogeneous FP requires 8 iterations and 14 seconds, while the fast FP just requires 2 iterations and 3.5 seconds. Thus, the fast FP can further improve upon the nonhomogeneous FP significantly, thanks to Nesterov's acceleration strategy.

V. Conclusion

This paper considers the ISAC beamforming optimization for a multi-cell massive MIMO network. Based on the FP technique, we first extend the classic WMMSE algorithm from the communications case to the ISAC case. This approach works but at a high cost of computation because the extended WMMSE requires inverting large matrices. We then develop the non-homogeneous bound to eliminate the large matrix inversion from the extended WMMSE (which can be also viewed as an FP based method). As a result, the per-iteration complexity can be significantly reduced. Furthermore, we show that the above FP based method is connected to gradient projection, and then propose using Nesterov's extrapolation scheme to render the beamforming algorithm converge much more rapidly.

References

[1] Y. Chen, Y. Feng, X. Li, L. Zhao, and K. Shen, "Fast fractional programming for multi-cell integrated sensing and communications," *arXiv preprint arXiv:2406.10910*, 2024.

[2] X. Liu, T. Huang, N. Shlezinger, Y. Liu, J. Zhou, and Y. C. Eldar, "Joint transmit beamforming for multiuser MIMO communications and MIMO radar," *IEEE Trans. Signal Process.*, vol. 68, pp. 3929–3944, June 2020.

[3] X. Li, F. Liu, Z. Zhou, G. Zhu, S. Wang, K. Huang, and Y. Gong, "Integrated sensing, communication, and computation over-the-air: MIMO beamforming design," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5383–5398, 2023.

[4] X. Wang, Z. Fei, J. A. Zhang, and J. Xu, "Partially-connected hybrid beamforming design for integrated sensing and communication systems," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6648–6660, Oct. 2022.

[5] Y. Guo, Y. Liu, Q. Wu, X. Li, and Q. Shi, "Joint beamforming and power allocation for RIS aided full-duplex integrated sensing and uplink communication system," *IEEE Trans. Wireless Commun.*, Oct. 2023.

[6] Z. Chen, J. Wang, Z. Tian, M. Wang, Y. Jia, and T. Q. Quek, "Joint rate splitting and beamforming design for RSMA-RIS-assisted ISAC system," *IEEE Wireless Commun. Lett.*, vol. 13, no. 1, pp. 173–177, Jan. 2023.

[7] K. Shen and W. Yu, "Fractional programming for communication systems—Part II: Uplink scheduling via matching," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2631–2644, Mar. 2018.

[8] K. Shen, W. Yu, L. Zhao, and D. P. Palomar, "Optimization of MIMO device-to-device networks via matrix fractional programming: A minorization–maximization approach," *IEEE/ACM Trans. Netw.*, vol. 27, no. 5, pp. 2164–2177, Oct. 2019.

[9] J. Zou, S. Sun, C. Masouros, Y. Cui, Y.-F. Liu, and D. W. K. Ng, "Energy-efficient beamforming design for integrated sensing and communications systems," *IEEE Trans. Commun.*, Feb. 2024.

[10] C. Deng, X. Fang, and X. Wang, "Beamforming design and trajectory optimization for UAV-empowered adaptable integrated sensing and communication," *IEEE Trans. Wireless Commun.*, vol. 22, no. 11, pp. 8512–8526, Nov. 2023.

[11] S. Fortunati, L. Sanguinetti, F. Gini, M. S. Greco, and B. Himed, "Massive MIMO radar for target detection," *IEEE Trans. Signal Process.*, vol. 68, pp. 859–871, Jan. 2020.

[12] S. S. Christensen, R. Agarwal, E. D. Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[13] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted mmse approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sept. 2011.

[14] X. Zhao, S. Lu, Q. Shi, and Z.-Q. Luo, "Rethinking WMMSE: Can its complexity scale linearly with the number of bs antennas?" *IEEE Trans. Signal Process.*, vol. 71, pp. 433–446, Feb. 2023.

[15] Z. Zhang, Z. Zhao, and K. Shen, "Enhancing the efficiency of WMMSE and fp for beamforming by minorization-maximization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, June 2023.

[16] K. Shen, Z. Zhao, Y. Chen, Z. Zhang, and H. V. Cheng, "Accelerating quadratic transform and WMMSE," *IEEE J. Sel. Areas Commun.*, 2024, to be published.

[17] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Aug. 2016.

[18] N. Su, F. Liu, and C. Masouros, "Sensing-assisted eavesdropper estimation: An ISAC breakthrough in physical layer security," *IEEE Trans. Wireless Commun.*, Aug. 2023.

[19] Z. Wang, X. Mu, and Y. Liu, "STARS enabled integrated sensing and communications," *IEEE Trans. Wireless Commun.*, vol. 22, no. 10, pp. 6750–6765, Oct. 2023.

[20] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Hoboken, NJ, USA: Wiley, 2002.

[21] Y. Nesterov, *Lectures on Convex Optimization (Second Edition)*. Springer, 2018.

[22] Z. Ren, Y. Peng, X. Song, Y. Fang, L. Qiu, L. Liu, D. W. K. Ng, and J. Xu, "Fundamental CRB-rate tradeoff in multi-antenna ISAC systems with information multicasting and multi-target sensing," *IEEE Trans. Wireless Commun.*, Sept. 2023.