

Co-Designing Statistical MIMO Radar and In-band Full-Duplex Multi-User MIMO Communications – Part II: Joint Precoder, Radar Code, and Receive Filters Design

Jiawei Liu^a, Kumar Vijay Mishra^b and Mohammad Saquib^a

^aThe University of Texas at Dallas, Richardson, TX 75080, USA

^bUnited States DEVCOM Army Research Laboratory, Adelphi, MD 20783, USA

ABSTRACT

We address the challenge of spectral sharing between a statistical multiple-input multiple-output (MIMO) radar and an in-band full-duplex (IBFD) multi-user MIMO (MU-MIMO) communications system operating simultaneously in the same frequency band. Existing research on joint MIMO-radar-MIMO-communications (MRMC) systems has limitations, such as focusing on colocated MIMO radars, half-duplex MIMO communications, single-user scenarios, neglecting practical constraints, or employing separate transmit/receive units for MRMC coexistence. This paper, along with companion papers (Part I and III), proposes a comprehensive MRMC framework that addresses all these challenges. In the previous companion paper (Part I), we presented signal processing techniques for a distributed IBFD MRMC system. In this paper, we introduce joint design of statistical MIMO radar codes, uplink/downlink precoders, and corresponding receive filters using a novel metric called compounded-and-weighted sum mutual information. To solve the resulting highly non-convex problem, we employ a combination of block coordinate descent (BCD) and alternating projection methods. Numerical experiments show convergence of our algorithm, mitigation of uplink interference, and stable data rates under varying noise levels, channel estimate imperfections, and self-interference. The subsequent companion paper (Part III) extends the discussion to multiple targets and evaluates the tracking performance of our MRMC system.

1. Introduction

The increasing congestion of the electromagnetic spectrum in recent years has presented significant challenges in the design of radar and communications systems operating within the same frequency bands [1]. While radar systems necessitate substantial transmit signal bandwidths for high-resolution target detection [2], wireless cellular networks require access to a broad spectrum to support high data rates [3, 4]. In response to the exponential growth of mobile data traffic, network operators globally have turned to higher frequency spectra to accommodate the surge in data usage [1]. Additionally, advancements in wireless communications and the continuous increase in carrier frequencies have prompted spectrum regulators such the Federal Communications Commission (FCC) and International Telecommunications Union (ITU) to grant civilian communications systems access to frequency bands traditionally reserved for radar and sensing applications. This policy shift has initiated a trend of coexistence and convergence between radar and communications functions [1].

The literature [1] suggests two possible approaches toward joint radar-communications. In the coexistence approach, the radar and communications systems operate as separate entities within the same spectrum using different waveforms [5, 6]. In the co-design paradigm, the two systems are integrated into a single hardware platform, and a common waveform is employed at either the transmitter (Tx), receiver (Rx), or both [7, 8]. The effectiveness of these spectrum-sharing solutions depends on the level of cooperation between the radar and communications systems. A hybrid approach of *spectral cooperation* has also been suggested, wherein some information exchange may take place between radar and communications systems [9, 10]. In this paper, we focus on spectral co-design aspects.

The aforementioned approaches do not readily extend to multiple-input multiple-output (MIMO) configurations, which employ multiple antennas at the transmitter and receiver to achieve high spectrum efficiency [3, 11]. MIMO

* K. V. M. acknowledges support from the National Academies of Sciences, Engineering, and Medicine via the Army Research Laboratory Harry Diamond Distinguished Fellowship. The following co-authors are a member of EURASIP: Kumar Vijay Mishra.

ORCID(s): 0000-0002-5386-609X (K.V. Mishra)

configurations enhance communications capacity, provide spatial diversity, and exploit multipath propagation [3]. Similarly, MIMO radars offer advantages over equivalent phased array radars, such as higher angular resolution with fewer antennas, spatial diversity, and improved parameter identification by leveraging waveform diversity [12]. In a colocated MIMO radar [13], the radar cross-section remains the same for closely-spaced antennas. In contrast, in a widely distributed or statistical MIMO radar, the antennas are sufficiently separated from each other, causing the same target to exhibit different radar cross-sections to each Tx-Rx pair. This spatial diversity is advantageous for detecting targets with small backscatters and low speed [14].

The increased degrees-of-freedom (DoFs), aperture sharing, and higher-dimensional optimization further complicate spectrum sharing in a joint MIMO-radar-MIMO-communications (MRMC) system [15, 16]. MRMC processing techniques include employing orthogonal transmit waveforms [17] and receiver interference cancellation [18]; see, e.g., [19] for a survey on MRMC solutions. Prior MRMC literature primarily focused on single-user MIMO communications and colocated MIMO radars. Co-design with statistical MIMO radar remains relatively unexamined in these prior works. In the previous companion paper (Part I) [19], we proposed spectral co-design of statistical MIMO radar with in-band full-duplex (IBFD) multi-user (MU) MIMO communications. The IBFD technology has been recently explored for joint radar-communications systems to facilitate communications transmission while also receiving the target echoes [20].

The performance metrics to design radar and communications systems are not identical because of different system goals [1]. As a result, recent works [15, 16] have suggested mutual information (MI) as a common metric for joint radar and communications systems. Our previous companion paper (Part I) [19] proposed MI-inspired novel *compounded-and-weighted sum MI* (CWSM) for the MIMO radar and IBFD MU-MIMO communications co-design problem. In that paper, we described the receive signal processing for a co-designed distributed MRMC system but did not develop an algorithm to solve the design problem. In this paper, unlike many prior works that focus solely on one specific system goal and often in isolation with other processing modules, we propose using CWSM to jointly design the UL/DL precoders, MIMO radar waveform matrix, and linear receive filters (LRFs) for both systems. Our co-design also accounts for several practical constraints, including the maximum UL/DL transmit powers, the QoS of the UL/DL quantified by their respective minimum achievable rates, and the peak-to-average-power-ratio (PAR) of the MIMO radar waveform. It is common among communications literature to identify a UL/DL UE's QoS with its minimum achievable rate [21, 22]. Adopting low PAR waveforms is crucial for achieving energy- and cost-efficient RF front-ends [23]. We address the non-convex CWSM maximization problem's challenges subject to non-convex constraints, namely the QoS and the PAR constraints, by developing an alternating algorithm that incorporates both the block coordinate descent (BCD) and the alternating projection (AP) methods. The BCD-AP process breaks the original problem into less complex subproblems that we iteratively solve. Numerical experiments show a quick, monotonic convergence of our proposed algorithm. Preliminary results of this work appeared in our conference publication [24], where only communications design was considered, PAR constraint was excluded, and detailed theoretical guarantees were excluded.

The rest of the paper is organized as follows. In the next section, we summarize the system model following the details included in the previous companion paper (Part I) [19]. Then, we formulate the CWSM maximization problem in Section 3. We develop the BCD-AP MRMC procedure to solve the non-convex optimization problem iteratively in Section 4. We validate the proposed technique through numerical experiments in Section 5 before concluding in Section 6.

Throughout this paper, lowercase regular, lowercase boldface and uppercase boldface letters denote scalars, vectors and matrices, respectively. We use $I(\mathbf{X}; \mathbf{Y})$ and $H(\mathbf{X}|\mathbf{Y})$ to denote, MI and conditional entropy between two random variables $\mathbf{X}$ and $\mathbf{Y}$, respectively. The notations $\mathbf{Y}[k]$, $\mathbf{y}[k]$, and $y[k]$ denote the value of time-variant matrix $\mathbf{Y}$, vector $\mathbf{y}$ and scalar y at discrete-time index k , respectively; $\mathbf{1}_N$ is a vector of size N with all ones; $\mathbb{C}$ and $\mathbb{R}$ represent sets of complex and real numbers, respectively; a circularly symmetric complex Gaussian (CSCG) vector $\mathbf{q}$ with N elements and power spectral density $\mathcal{N}_0$ is $\mathbf{q} \sim \mathcal{CN}(0, \mathcal{N}_0 \mathbf{I}_N)$; $(\cdot)^*$ is the solution of the optimization problem; $\mathbb{E}[\cdot]$ is the statistical expectation; $\text{Tr}\{\mathbf{R}\}$, $\mathbf{R}^\top$, $\mathbf{R}^\dagger$, $\mathbf{R}^*$, $|\mathbf{R}|$, $\mathbf{R} \geq \mathbf{0}$, and $\mathbf{R}(m, n)$ are the trace, transpose, Hermitian transpose, element-wise complex conjugate, determinant, positive semi-definiteness and (m, n) -th entry of matrix $\mathbf{R}$, respectively; set $\mathbb{Z}_+(L)$ denotes $\{1, \dots, L\}$; $\mathbf{x} \geq \mathbf{y}$ denotes component-wise inequality between vectors $\mathbf{x}$ and $\mathbf{y}$; x^+ represents $\max(x, 0)$; $x^{(t)}(\cdot)$ is the t -th iterate of an iterative function $x(\cdot)$; $\inf(\cdot)$ is the infimum of its argument; $\odot$ denotes the Hadamard product; and $\oplus$ is the direct sum.

2. Spectral Co-Design System Model

The system model for the statistical MIMO radar signal model used in this paper closely follows that detailed in the previous companion paper (Part I) [19] and, hence, we only summarize the key aspects below.

Consider a two-dimensional (2-D) $(x-y)$ Cartesian plane on which the M_r Tx and N_r Rx of a statistical MIMO radar, the BS, I UL UEs, and J DL UEs of the IBFD MU-MIMO communications system are located at the coordinates (x_{m_r}, y_{m_r}) , (x_{n_r}, y_{n_r}) , (x_B, y_B) , $(x_{UL,i}, y_{UL,i})$, and $(x_{DL,j}, y_{DL,j})$, respectively, for all $m_r \in \mathbb{Z}_+(M_r)$, $n_r \in \mathbb{Z}_+(N_r)$, $i \in \mathbb{Z}_+(I)$, and $j \in \mathbb{Z}_+(J)$. The statistical MIMO radar operates within the same transmit spectrum as an IBFD MU-MIMO communications system. Here, the radar aims to detect a target moving within the cellular coverage of the BS. The communications system serves the UL/DL UEs with desired achievable rates in the presence of the radar echoes.

2.1. Transmit Signal

Each radar Tx emits a train of K pulses at a uniform pulse repetition interval (PRI) T_r or *fast-time*; the total duration KT_r is the *coherent processing interval* (CPI) or *slow-time* and K is chosen to avoid range migration during the CPI [2]. At the same time, the BS and each UL UE continuously transmit DL and UL symbols, respectively. The radar pulse width is $T_p = T_r/N$, where N is the number of sampled range bins or cells in a PRI. The UL/DL frame duration T_f and the UL/DL symbol duration T_s equal radar PRI and radar pulse width T_p , respectively; i.e., $T_f = T_r$ and $T_s = T_p$. This implies that the number of UL/DL frames transmitted in the scheduling window is also K and the number of UL/DL symbols per frame is $T_f/T_s = N$. The k -th communications frame is transmitted at a duration of GT_p , $G \in \mathbb{Z}(N-1)$, before the k -th radar PRI.

2.1.1. Statistical MIMO Radar

Denote the *narrowband* transmit pulse of the m_r -th radar Tx by $\phi_{m_r}(t)$. The waveforms from all Tx form the waveform vector

$$\boldsymbol{\phi}(t) = [\phi_1(t), \dots, \phi_{M_r}(t)]^\top \in \mathbb{C}^{M_r}, \quad (1)$$

and satisfy the orthonormality $\int_{T_p} \boldsymbol{\phi}(t) \boldsymbol{\phi}^\dagger(t) dt = \mathbf{I}_{M_r}$. The radar code to modulate the pulse emitted by the m_r Tx in the k -th PRI is $a_{m_r,k}$. During the observation window $t \in [0, KT_r + GT_p]$, the m_r -th Tx emits the pulse train

$$s_{m_r}(t) = \sum_{k=0}^{K-1} a_{m_r,k} \phi_{m_r}(t - kT_r - GT_p), \quad (2)$$

where the support of $\phi_{m_r}(t)$ is $[0, T_r)$ and, without loss of generality, $\phi_{m_r}(t) = \sqrt{1/T_p} e^{j2\pi \frac{m_r}{T_p} t}$ for $m_r \in \mathbb{Z}_+(M_r)$ for $t \in [0, T_r)$. Define the radar code vector transmitted during the k -th PRI as $\mathbf{a}[k] = [a_{1,k}, \dots, a_{M_r,k}]^\top \in \mathbb{C}^{M_r}$ so that the MIMO radar code matrix is

$$\mathbf{A} = [\mathbf{a}^\top[1]; \dots; \mathbf{a}^\top[K]] = [\mathbf{a}_1, \dots, \mathbf{a}_{M_r}] \in \mathbb{C}^{K \times M_r}. \quad (3)$$

where $\mathbf{a}_{m_r} \in \mathbb{C}^K$ is the code of the m_r -th TX over all PRIs. The combined transmit signal vector is

$$\mathbf{s}(t) = [s_1(t), \dots, s_{M_r}(t)]^\top \in \mathbb{C}^{M_r}. \quad (4)$$

2.1.2. IBFD MU-MIMO Communications

The BS and UEs operate in the FD and HD modes, respectively. During the observation window, the BS receives data frames from the I UL UEs; concurrently, the J DL UEs operating in the same band download data frames from the BS. The BS is equipped with M_c transmit and N_c receive antennas. The i -th UL UE and j -th DL UE employ N_i^u and N_j^d transceive antennas, respectively. To achieve the maximum capacities of the UL and DL channels, number of BS Tx and Rx antennas are $M_c \geq \sum_{j=1}^J N_j^d$ and $N_c \geq \sum_{i=1}^I N_i^u$, respectively [3]. A total of $D_i^u \leq N_i^u$ and

$D_j^d \leq N_j^d$ unit-energy data streams are used by i -th UL UE and j -th DL UE, respectively. The symbol vectors sent by the i -th UL UE toward the BS and by the BS toward the j -th DL UE in the l -th symbol period of the k -th frame are $\mathbf{d}_{u,i}[k, l] \in \mathbb{C}^{D_i^u}$ and $\mathbf{d}_{d,j}[k, l] \in \mathbb{C}^{D_j^d}$, respectively; these are independent and identically distributed (i.i.d.) with $\mathbb{E}[\mathbf{d}_{d,j}\mathbf{d}_{d,j}^\dagger[k, l]] = \mathbb{E}[\mathbf{d}_{u,i}\mathbf{d}_{u,i}^\dagger[k, l]] = \mathbf{I}$ for $i \in \mathbb{Z}_+\{I\}$, $k \in \mathbb{Z}_+\{K\}$, and $l \in \mathbb{Z}_+\{N\}$.

Denote the precoders for the i -th UL UE and the j -th DL UE at the k -th frame as $\mathbf{P}_{u,i}[k] \in \mathbb{C}^{N_i^u \times D_i^u}$ and $\mathbf{P}_{d,j}[k] \in \mathbb{C}^{M_c \times D_j^d}$, respectively. The precoded transmit signal vectors for the i -th UL UE and j -th DL UE become

$$\mathbf{s}_{u,i}[k, l] = \mathbf{P}_{u,i}[k]\mathbf{d}_{u,i}[k, l], \quad (5)$$

and

$$\mathbf{s}_{d,j}[k, l] = \mathbf{P}_{d,j}[k]\mathbf{d}_{d,j}[k, l], \quad (6)$$

respectively. The total DL symbol vector broadcast by the BS in the same symbol period is

$$\mathbf{s}_B[k, l] = \sum_{j=1}^J \mathbf{s}_{d,j}[k, l]. \quad (7)$$

The transmit pulse shaping function used by the IBFD communications is $p_T(t)$. The transmit signals of i -th UL UE and BS are

$$\mathbf{x}_{u,i}(t) = \sum_{k=0}^{K-1} \sum_{l=0}^{N-1} \mathbf{s}_{u,i}[k, l] p_T(t - (kN + l)T_p), \quad (8)$$

and

$$\mathbf{x}_B(t) = \sum_{k=0}^{K-1} \sum_{l=0}^{N-1} \mathbf{s}_B[k, l] p_T(t - (kN + l)T_p). \quad (9)$$

2.2. Statistical MIMO radar receiver

The radar and DL signals are reflected off a single target and the combined echo is received at the n_r -th radar Rx as $\mathbf{y}_{t,n_r}$. It is overlaid with clutter echoes $\mathbf{y}_{c,n_r}$, directly received IBFD MU-MIMO DL signal $\mathbf{y}_{Bm,n_r}$, and interference from the UL $\mathbf{y}_{u,n_r}$. With the CSCG noise vector at the n_r -th radar Rx by $\mathbf{z}_{r,n_r} \in \mathcal{CN}(\mathbf{0}, \sigma_{r,n_r}^2 \mathbf{I}_K)$, the composite receive signal model at the range cell under test (CUT) of the n_r -th radar Rx is

$$\mathbf{y}_{r,n_r} = \mathbf{y}_{t,n_r} + \underbrace{\mathbf{y}_{c,n_r} + \mathbf{y}_{Bm,n_r} + \mathbf{y}_{u,n_r} + \mathbf{z}_{r,n_r}}_{=\mathbf{y}_{r,n_r}^{\text{in}}}, \quad (10)$$

where $\mathbf{y}_{r,n_r}^{\text{in}}$ denotes the interference-plus-noise component of $\mathbf{y}_{r,n_r}$. The covariance matrix (CM) of $\mathbf{y}_{r,n_r}$ is [19]

$$\mathbf{R}_{r,n_r} = \mathbf{R}_{t,n_r} + \mathbf{R}_{r,n_r}^{\text{in}}, \quad (11)$$

with the CM of $\mathbf{y}_{r,n_r}^{\text{in}}$ given by

$$\mathbf{R}_{r,n_r}^{\text{in}} \triangleq \mathbf{R}_{c,n_r} + \mathbf{R}_{Bm,n_r} + \mathbf{R}_{u,n_r} + \sigma_{r,n_r}^2 \mathbf{I}_K. \quad (12)$$

Combining the received signals from N_r radar Rxs yields

$$\mathbf{y}_r = \mathbf{y}_{tr} + \mathbf{y}_r^{\text{in}} = \begin{bmatrix} \mathbf{y}_{r,1}^\top; \dots; \mathbf{y}_{r,N_r}^\top \end{bmatrix}^\top \in \mathbb{C}^{KN_r}, \quad (13)$$

where

$$\mathbf{y}_{\text{tr}} = \left[\mathbf{y}_{\text{t},1}^\top; \dots; \mathbf{y}_{\text{t},N_r}^\top \right]^\top, \quad (14)$$

and

$$\mathbf{y}_{\text{r}}^{\text{in}} \triangleq \mathbf{y}_{\text{cr}} + \mathbf{y}_{\text{Bmr}} + \mathbf{y}_{\text{Ur}} + \mathbf{z}_{\text{r}} = \left[\left(\mathbf{y}_{\text{r},1}^{\text{in}} \right)^\top; \dots; \left(\mathbf{y}_{\text{r},N_r}^{\text{in}} \right)^\top \right]^\top, \quad (15)$$

whose CM is

$$\mathbf{R}_{\text{r}}^{\text{in}} = \bigoplus_{n_r=1}^{N_r} \mathbf{R}_{\text{r},n_r}^{\text{in}}. \quad (16)$$

2.3. IBFD MU-MIMO Communications Receiver

Within the observation window, J DL UEs and the BS receive both IBFD communications signals and radar probing signals. For l -th symbol period and k -th frame, denote the CSCG noise vectors measured respectively at the BS Rx and the j -th DL UE as $\mathbf{z}_{\text{B}}[k, l] \sim \mathcal{CN}(0, \sigma_{\text{B}}^2 \mathbf{I}_{M_{\text{c}}})$ and $\mathbf{z}_{\text{d},j}[k, l] \sim \mathcal{CN}(0, \sigma_{\text{d},j}^2 \mathbf{I}_{N_j^{\text{d}}})$, i.i.d in k and l . Then, the signal received at the BS Rx to decode $\mathbf{s}_{\text{u},i}[k, l]$ and the composite signal received by the j -th DL UE are, respectively,

$$\mathbf{y}_{\text{u},i}[k, l] = \mathbf{y}_{\text{i},\text{B}}[k, l] + \mathbf{y}_{\text{um},i}[k, l] + \mathbf{y}_{\text{BB}}[k, l] + \mathbf{y}_{\text{rB}}[k, l] + \mathbf{z}_{\text{B}}[k, l], \quad (17)$$

$$\text{and } \mathbf{y}_{\text{d},j}[k, l] = \mathbf{y}_{\text{B},j}[k, l] + \mathbf{y}_{\text{dm},j}[k, l] + \mathbf{y}_{\text{u},j}[k, l] + \mathbf{y}_{\text{r},j}[k, l] + \mathbf{z}_{\text{d},j}[k, l], \quad (18)$$

where $\mathbf{y}_{\text{i},\text{B}}[k, l]$ is the signal from i -th UL UE to the BS; $\mathbf{y}_{\text{um},i}[k, l]$ is the multi-user interference (MUI) from other UL UEs; $\mathbf{y}_{\text{BB}}[k, l]$ is the self-interference due to FD transmission; $\mathbf{y}_{\text{rB}}[k, l]$ ($\mathbf{y}_{\text{r},j}[k, l]$) is the radar signal received at the BS (j -th DL UE); $\mathbf{y}_{\text{B},j}[k, l]$ is the signal from the BS aimed for the j -th DL UE; $\mathbf{y}_{\text{dm},j}[k, l]$ is the MUI from other DL UEs; and $\mathbf{y}_{\text{u},j}[k, l]$ is the UL interference at the j -th DL UE. The CMs of $\mathbf{y}_{\text{u},i}[k, l]$ and $\mathbf{y}_{\text{d},j}[k, l]$ are, respectively,

$$\mathbf{R}_{\text{u},i}[k, l] = \mathbf{R}_{\text{i},\text{B}}[k, l] + \mathbf{R}_{\text{u},i}^{\text{in}}[k, l], \quad (19)$$

and

$$\mathbf{R}_{\text{d},j}[k, l] = \mathbf{R}_{\text{B},j}[k, l] + \mathbf{R}_{\text{d},j}^{\text{in}}[k, l], \quad (20)$$

where

$$\mathbf{R}_{\text{u},i}^{\text{in}}[k, l] = \mathbf{R}_{\text{um},i}[k, l] + \mathbf{R}_{\text{BB}}[k, l] + \mathbf{R}_{\text{rB}}[k, l] + \sigma_{\text{B}}^2 \mathbf{I}_{M_{\text{c}}}, \quad (21)$$

and

$$\mathbf{R}_{\text{d},j}^{\text{in}}[k, l] = \mathbf{R}_{\text{dm},j}[k, l] + \mathbf{R}_{\text{u},j}[k, l] + \mathbf{R}_{\text{r},j}[k, l] + \sigma_j^2 \mathbf{I}_{N_j^{\text{d}}}, \quad (22)$$

denote the interference-plus-noise CMs associated with (18) and (17), respectively. The precoders of IBFD communications are based on the n_{rB} -th symbol period of K UL frames and the n_{rd} -th symbol period of K DL frames, where n_{rB} and n_{rd} are the symbol indices of UL and DL, respectively. Figure 1 illustrates the composite receive signals of BS Rx, j -th DL UE, and n_{r} -th radar Rx.

3. CWSM Maximization

We now define the LRFs for the MIMO radar and the IBFD MU-MIMO communications system Rxs before introducing the MI-based co-design metric CWSM. Denote the LRF at the n_{r} -th radar as $\mathbf{U}_{\text{r},n_{\text{r}}} = \left[\mathbf{u}_{\text{r},n_{\text{r}}}[0], \dots, \mathbf{u}_{\text{r},n_{\text{r}}}[K-1] \right] \in \mathbb{C}^{KM \times K}$. This LRF's output is

$$\tilde{\mathbf{y}}_{\text{r},n_{\text{r}}} = \tilde{\mathbf{y}}_{\text{t},n_{\text{r}}} + \tilde{\mathbf{y}}_{\text{r},n_{\text{r}}}^{\text{in}} = \mathbf{U}_{\text{r},n_{\text{r}}} \mathbf{S}_{\text{t},n_{\text{r}}} \mathbf{h}_{\text{t},n_{\text{r}}} + \mathbf{U}_{\text{r},n_{\text{r}}} \mathbf{y}_{\text{r},n_{\text{r}}}^{\text{in}} \quad (23)$$

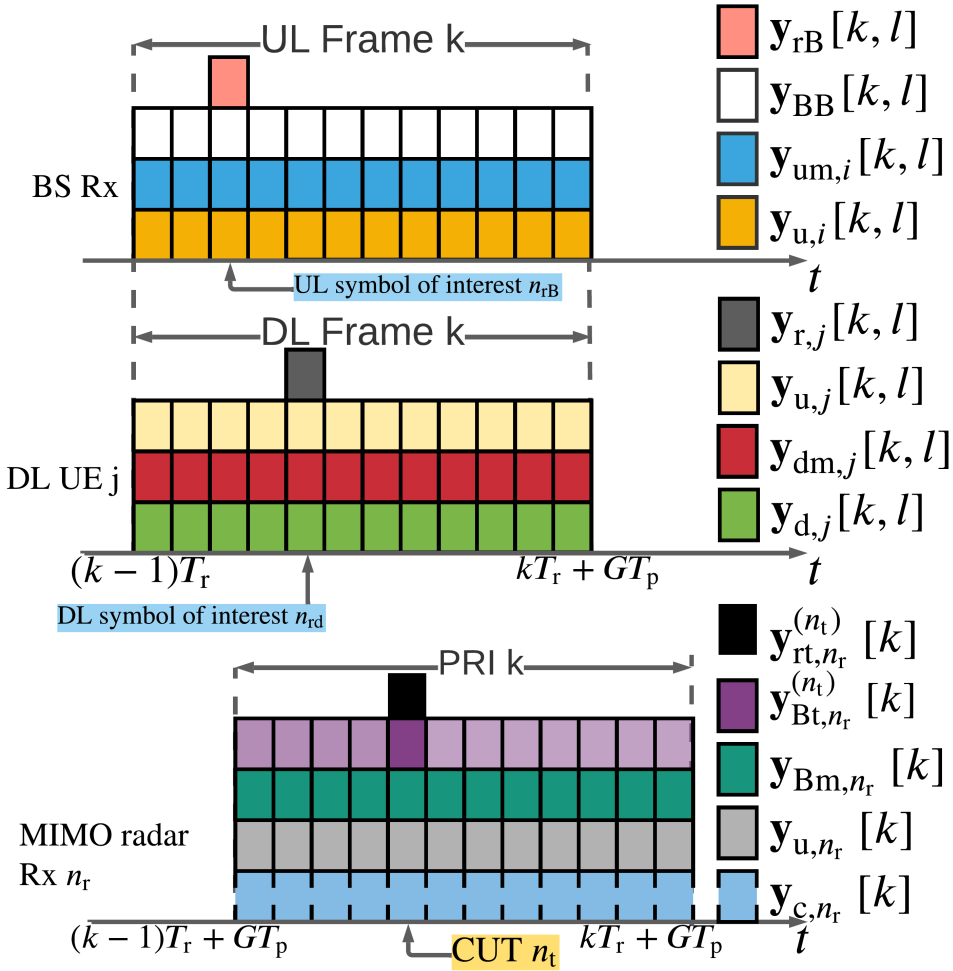


Figure 1: The overlaid receive signal timing diagram during k -th radar PRI and k -th communications frame in the observation window; noise trails have been excluded. The purple bin with more opacity indicates the DL signal reflected from the target and observed in the radar CUT, i.e., $\mathbf{y}_{\text{Bt},n_r}^{(n_t)}[k]$.

where $\mathbf{h}_{t,n_r} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_{t,n_r})$ contains the target information and $\tilde{\mathbf{y}}_{r,n_r} \sim \mathcal{CN}(\mathbf{0}, \mathbf{U}_{r,n_r}(\mathbf{R}_{t,n_r} + \mathbf{R}_{r,n_r}^{\text{in}})\mathbf{U}_{r,n_r}^\dagger)$.

Using the chain rule, the MI between $\tilde{\mathbf{y}}_{r,n_r}$ and $\mathbf{h}_{t,n_r}$ is [25, 26]

$$\begin{aligned} I_{r,n_r} &\triangleq I(\tilde{\mathbf{y}}_{r,n_r}; \mathbf{h}_{t,n_r}) = H(\tilde{\mathbf{y}}_{r,n_r}) - H(\tilde{\mathbf{y}}_{r,n_r} | \mathbf{h}_{t,n_r}) \\ &= H(\tilde{\mathbf{y}}_{r,n_r}) - H(\tilde{\mathbf{y}}_{t,n_r} | \mathbf{h}_{t,n_r}) - H(\tilde{\mathbf{y}}_{r,n_r}^{\text{in}} | \mathbf{h}_{t,n_r}) \\ &= H(\tilde{\mathbf{y}}_{r,n_r}) - H(\tilde{\mathbf{y}}_{r,n_r}^{\text{in}}), \end{aligned} \quad (24)$$

where $H(\tilde{\mathbf{y}}_{t,n_r} | \mathbf{h}_{t,n_r})$ vanishes because $\tilde{\mathbf{y}}_{t,n_r}$ depends on $\mathbf{h}_{t,n_r}$; and $H(\tilde{\mathbf{y}}_{r,n_r}^{\text{in}} | \mathbf{h}_{t,n_r})$ reduces to $H(\tilde{\mathbf{y}}_{r,n_r}^{\text{in}})$ because $\tilde{\mathbf{y}}_{r,n_r}^{\text{in}}$ and $\mathbf{h}_{t,n_r}$ are mutually independent. The conditional differential entropy with the Gaussian noise [25] leads to

$$H(\tilde{\mathbf{y}}_{r,n_r} | \mathbf{A}) = \varphi + \log |\mathbf{U}_{r,n_r} \mathbf{R}_{r,n_r} \mathbf{U}_{r,n_r}^\dagger| \quad (25)$$

and

$$H(\tilde{\mathbf{y}}_{r,n_r}^{\text{in}} | \mathbf{A}) = \varrho + \log \left| \mathbf{U}_{r,n_r} \mathbf{R}_{r,n_r}^{\text{in}} \mathbf{U}_{r,n_r}^\dagger \right|, \quad (26)$$

where the constant $\varrho = K \log(\pi) + K$. This gives

$$\begin{aligned} I_{r,n_r} &= \log \frac{\left| \mathbf{U}_{r,n_r} \left(\mathbf{R}_{t,n_r} + \mathbf{R}_{r,n_r}^{\text{in}} \right) \mathbf{U}_{r,n_r}^\dagger \right|}{\left| \mathbf{U}_{r,n_r} \mathbf{R}_{r,n_r}^{\text{in}} \mathbf{U}_{r,n_r}^\dagger \right|} \\ &= \log \left| \mathbf{I} + \mathbf{U}_{r,n_r} \mathbf{R}_{t,n_r} \mathbf{U}_{r,n_r}^\dagger \left(\mathbf{U}_{r,n_r} \mathbf{R}_{r,n_r}^{\text{in}} \mathbf{U}_{r,n_r}^\dagger \right)^{-1} \right|. \end{aligned} \quad (27)$$

The LRFs deployed at the BS to decode the i -th UL UE and j -th DL UE during k -th frame of the observation window are $\mathbf{U}_{u,i}[k] \in \mathbb{C}^{D_l^i \times N_c}$ and $\mathbf{U}_{d,j}[k] \in \mathbb{C}^{D_j^d \times N_j^d}$, respectively. The outputs of $\mathbf{U}_{u,i}[k]$ and $\mathbf{U}_{d,j}[k]$ are $\tilde{\mathbf{y}}_{u,i}[k, l] = \mathbf{U}_{u,i}[k] \mathbf{y}_{u,i}[k, l]$ and $\tilde{\mathbf{y}}_{d,j}[k, l] = \mathbf{U}_{d,j}[k] \mathbf{y}_{d,j}[k, l]$; these signals follow the distributions $\mathcal{CN}(\mathbf{0}, \mathbf{U}_{u,i}[k] \mathbf{R}_{u,i}[k, l] \mathbf{U}_{u,i}^\dagger[k])$ and $\mathcal{CN}(\mathbf{0}, \mathbf{U}_{d,j}[k] \mathbf{R}_{d,j}[k, l] \mathbf{U}_{d,j}^\dagger[k])$, respectively.

Using the common assumption of Gaussianity on symbol vectors, i.e. $\mathbf{s}_{u,i}[k, l] \sim \mathcal{CN}(\mathbf{0}, \mathbf{P}_{u,i}[k] \mathbf{P}_{u,i}^\dagger[k])$ and $\mathbf{s}_{d,j}[k, l] \sim \mathcal{CN}(\mathbf{0}, \mathbf{P}_{d,j}[k] \mathbf{P}_{d,j}^\dagger[k])$, the MIs between $\tilde{\mathbf{y}}_{u,i}[k, l]$ and $\mathbf{s}_{u,i}[k, l]$ as well as $\tilde{\mathbf{y}}_{d,j}[k, l]$ and $\mathbf{s}_{d,j}[k, l]$ are

$$I_i^u[k, l] \triangleq I(\mathbf{s}_{u,i}[k, l]; \tilde{\mathbf{y}}_{u,i}[k, l]) = \log \left| \mathbf{I} + \mathbf{U}_{u,i}[k] \mathbf{R}_{i,B}[k, l] \mathbf{U}_{u,i}^\dagger[k] \left(\mathbf{U}_{u,i}[k] \mathbf{R}_{u,i}^{\text{in}}[k, l] \mathbf{U}_{u,i}^\dagger[k] \right)^{-1} \right|, \quad (28)$$

and

$$I_j^d[k, l] \triangleq I(\mathbf{s}_{B,j}[k, l]; \tilde{\mathbf{y}}_{d,j}[k, l]) = \log \left| \mathbf{I} + \mathbf{U}_{d,j}[k] \mathbf{R}_{B,j}[k, l] \mathbf{U}_{d,j}^\dagger[k] \left(\mathbf{U}_{d,j}[k] \mathbf{R}_{d,j}^{\text{in}}[k, l] \mathbf{U}_{d,j}^\dagger[k] \right)^{-1} \right|, \quad (29)$$

respectively. We evaluate FD communications based on the symbols-of-interest, i.e., $l = n_{\text{rB}}$ for each UL frame and $l = n_{\text{rd}}$ for each DL frame in the observation window. The metric CWSM is a weighted sum of communications' MIs related to the symbol periods of interest¹ and I_{r,n_r} , i.e.,

$$I_{\text{CWSM}} = \sum_{n_r=1}^{N_r} \alpha_{n_r}^r I_{n_r}^r + \sum_{k=0}^{K-1} \left[\sum_{i=1}^I \alpha_i^u I_i^u[k] + \sum_{j=1}^J \alpha_j^d I_j^d[k] \right], \quad (30)$$

where $\alpha_{n_r}^r$, α_i^u , and α_j^d are pre-defined weights assigned to the MIMO n_r -th radar Rx, i -th UL UE and j -th DL UE, respectively, for all n_r, i , and j ; the weights are determined by the system priority and specific applications. For example, for FD communications, the weights are based on available buffer capacities of BS and UEs [27, 28]. For the joint radar-communications, one can assign larger (smaller) weights to $\alpha_{n_r}^r$ with the presence (absence) of targets [29].

Denote the sets of the precoders and the LRFs as $\{\mathbf{P}\} \triangleq \{\mathbf{P}_{u,i}[k], \mathbf{P}_{d,j}[k] | i \in \mathbb{Z}_+\{I\}, j \in \mathbb{Z}_+\{J\}, k \in \mathbb{Z}_+\{K\}\}$ and $\{\mathbf{U}\} \triangleq \{\mathbf{U}_{u,i}[k], \mathbf{U}_{d,j}[k], \mathbf{U}_{r,n_r} | i \in \mathbb{Z}_+\{I\}, j \in \mathbb{Z}_+\{J\}, k \in \mathbb{Z}_+\{K\}, n_r \in \mathbb{Z}_+\{N_r\}\}$. The transmission powers that occurred to the BS and the i -th UL UE at the k -th frame are

$$P_d[k] = \sum_{j=1}^J P_{d,j}[k] = \sum_{j=1}^J \text{Tr} \left\{ \mathbf{P}_{d,j}[k] \mathbf{P}_{d,j}^\dagger[k] \right\}, \quad (31)$$

and

$$P_{u,i}[k] = \text{Tr} \left\{ \mathbf{P}_{u,i}[k] \mathbf{P}_{u,i}^\dagger[k] \right\}, \quad (32)$$

¹Hereafter, for simplicity, we drop symbol index $l = n_{\text{rB}}$ ($l = n_{\text{rd}}$) for UL (DL) related terms.

which are upper bounded by the maximum DL and UL powers P_B and P_U , respectively. The achievable rates for the i -th UL UE and the j -th DL UE in the k -th frame, $R_{u,i}[k]$ and $R_{d,j}[k]$ are lower bounded by the least acceptable achievable rates to quantify the QoS of the UL and DL, R_{UL} and R_{DL} , respectively. The CWSM optimization to jointly design precoders $\{\mathbf{P}\}$, radar code $\mathbf{A}$, and LRFs $\{\mathbf{U}\}$ is

$$\underset{\{\mathbf{P}\}, \{\mathbf{U}\}, \mathbf{A}}{\text{maximize}} \quad I_{\text{CWSM}}(\{\mathbf{U}\}, \{\mathbf{P}\}, \mathbf{A}) \quad (33a)$$

$$\text{subject to} \quad P_d[k] \leq P_B, \quad (33b)$$

$$P_{u,i}[k] \leq P_U, \quad (33c)$$

$$R_{u,i}[k] \geq R_{UL}, \quad (33d)$$

$$R_{d,j}[k] \geq R_{DL}, \quad (33e)$$

$$\|\mathbf{a}_{m_r}\|^2 = P_{r,m_r}, \quad (33f)$$

$$\frac{K \max_{k=1, \dots, K} |\mathbf{a}_{m_r}[k]|^2}{P_{r,m_r}} \leq \gamma_{m_r}, \quad \forall i, j, k, m_r, \quad (33g)$$

where constraints (33f) and (33g) are determined by the transmit power and PAR of the m_r -th MIMO radar Tx, respectively. Note that PAR constraint is applied column-wise to the code matrix $\mathbf{A}$ because the Tx's of a statistical MIMO radar are widely distributed. When $\gamma_{m_r} = 1$, PAR constraint is reduced to constant modulus constraint.

4. Joint Code-Precoder-Filter Design

Even without the non-convex constraints (33e)-(33g), (33) is non-convex because the objective function I_{CWSM} is not jointly concave over $\{\mathbf{P}\}$, $\{\mathbf{U}\}$, and $\mathbf{A}$, and therefore its global optima are generally intractable [30]. In general, such a problem is solved by alternately optimizing over one unknown variable at a time. When the number of variables is large, methods such as the BCD partition all optimization variables into, say, V small groups or blocks and optimize over each block, one at a time, while keeping other block variables fixed [31]. The net effect is that the problem is equivalently solved by iteratively solving less complex V subproblems. If there are only two blocks of variables, the BCD reduces to the classical alternating minimization method [31, 32].

It has been shown [33, 34] that the BCD converges globally to a stationary point for both convex and non-convex problems while methods such as Alternating Direction Method of Multipliers (ADMM) and Douglas-Rachford Splitting (DRS) achieve only linear convergence for strictly convex and some non-convex (e.g. multi-convex) problems. The stochastic gradient descent used to address saddle point problems has a slower convergence rate than BCD and offers only weak convergence for non-convex problems [34].

One can partition the block coordinate variables from (33) into three groups, i.e., $\{\mathbf{P}\}$, $\mathbf{A}$, and $\{\mathbf{U}\}$. At each iteration, we apply a *direct update* [31], i.e., maximize I_{CWSM} for all the block variables. Further, we update the block variables in a cyclic sequence because its global and local convergence has been well-established [31, 35] compared to other sequential update rules². In particular, we adopt the Gauss-Seidel BCD [31], which minimizes the objective function cyclically over each block while keeping the other blocks fixed.

A summary of our strategy is as follows. Note that $\{\mathbf{P}\}$ is subject to only communications-centric constraints (33b)-(33e) and $\mathbf{A}$ to both radar-centric PAR constraints (33f)-(33g) and communications-centric constraints (33b)-(33e). We denote the sets of feasible $\mathbf{A}$ determined by (33b)-(33e) and (33f)-(33g) as $\mathbb{A}_c$ and $\mathbb{A}_r$, respectively and the optimal solution for $\mathbf{A}$, i.e., $\mathbf{A}^*$ is thus in the intersection of $\mathbb{A}_c$ and $\mathbb{A}_r$, i.e., $\mathbf{A}^* \in \mathbb{A}_c \cap \mathbb{A}_r$. The AP method [36, 37] is appropriate to perform the search for $\mathbf{A}^*$. In the sequel of this section, we first transform the I_{CWSM} maximization problem, which is a weighted sum rate (WSR) problem, in (33) to a weighted minimum mean-squared-error (WMMSE) minimization problem without the PAR constraints in Section 4.1. It has been shown in [38, 28] that maximizing information-theoretic quantities via WMMSE for MIMO precoder design yields better results than geometric programming. Mapping a WSR problem to its corresponding WMMSE problem is also more computationally efficient than gradient-based (GB) approaches [28] because of its low per-iteration complexity, which is guaranteed to converge to at least a local optimum [38]. In Section 4.2, we then employ the BCD in our proposed WMMSE-MRMC algorithm to solve

²A possible alternative is the randomized BCD, where the series of iterates generated by BCD are divergent. However, cyclic BCD may still outperform the randomized BCD [33].

for the optimal $\mathbf{A}$ in $\mathbb{A}_c$, which we denote as $\mathbf{A}'$, and the optimal $\{\mathbf{P}^\star\}$. Then, we project each column of $\mathbf{A}'$ onto $\mathbb{A}_r$ in Section 4.3 using AP. The WMMSE-MRMC and AP procedures comprise the overall BCD-AP MRMC algorithm (Algorithm 4) and are repeated until convergence.

4.1. Relationship between WMMSE and MI

Consider (33) without PAR constraints,

$$\underset{\{\mathbf{P}\}, \{\mathbf{U}\}, \mathbf{A}}{\text{maximize}} \quad I_{\text{CWSM}}(\{\mathbf{U}\}, \{\mathbf{P}\}, \mathbf{A}) \quad \text{subject to (33b) – (33e)}. \quad (34)$$

To derive the WMMSE expressions regarding (34), we first define the mean squared error for the n_r -th radar Rx, i -th UL UE, and j -th DL UE as

$$\begin{aligned} \mathbf{E}_{r,n_r} &= \mathbb{E} \left[\left(\mathbf{h}_{t,n_r} - \mathbf{U}_{r,n_r} \mathbf{y}_{r,n_r} \right) \left(\mathbf{h}_{t,n_r} - \mathbf{U}_{r,n_r} \mathbf{y}_{r,n_r} \right)^\dagger \right] \\ &= \mathbf{\Sigma}_{t,n_r} - \mathbf{U}_{r,n_r} \mathbf{S}_{t,n_r} \mathbf{\Sigma}_{t,n_r}^\dagger - \mathbf{\Sigma}_{t,n_r}^\dagger \mathbf{S}_{t,n_r}^\dagger \mathbf{U}_{r,n_r}^\dagger \\ &\quad + \mathbf{U}_{r,n_r} \mathbf{R}_{r,n_r} \mathbf{U}_{r,n_r}^\dagger, \end{aligned} \quad (35)$$

$$\begin{aligned} \mathbf{E}_{u,i}[k] &= \mathbb{E} \left[\left(\mathbf{d}_{u,i}[k] - \mathbf{U}_{u,i}[k] \mathbf{y}_{u,i}[k] \right) \left(\mathbf{d}_{u,i}[k] - \mathbf{U}_{u,i}[k] \mathbf{y}_{u,i}[k] \right)^\dagger \right] \\ &= \mathbf{I} - \mathbf{U}_{u,i}[k] \mathbf{H}_{i,B} \mathbf{P}_{u,i}[k] - \mathbf{P}_{u,i}^\dagger[k] \mathbf{H}_{i,B}^\dagger \mathbf{U}_{u,i}^\dagger[k] \\ &\quad + \mathbf{U}_{u,i}[k] \mathbf{R}_{u,i}[k] \mathbf{U}_{u,i}^\dagger[k], \end{aligned} \quad (36)$$

$$\begin{aligned} \mathbf{E}_{d,j}[k] &= \mathbb{E} \left[\left(\mathbf{d}_{d,j}[k] - \mathbf{U}_{d,j}[k] \mathbf{y}_{d,j}[k] \right) \left(\mathbf{d}_{d,j}[k] - \mathbf{U}_{d,j}[k] \mathbf{y}_{d,j}[k] \right)^\dagger \right] \\ &= \mathbf{I} - \mathbf{U}_{d,j}[k] \mathbf{H}_{B,j} \mathbf{P}_{d,j}[k] - \mathbf{P}_{d,j}^\dagger[k] \mathbf{H}_{B,j}^\dagger \mathbf{U}_{d,j}^\dagger[k] \\ &\quad + \mathbf{U}_{d,j}[k] \mathbf{R}_{d,j}[k] \mathbf{U}_{d,j}^\dagger[k], \end{aligned} \quad (37)$$

where the expectations are taken w.r.t. $\mathbf{h}_{t,n_r}$, $\mathbf{d}_{u,i}[k]$, and $\mathbf{d}_{d,j}[k]$, respectively. Denote symmetric weight matrices associated with $\mathbf{E}_{r,n_r}$, $\mathbf{E}_{u,i}[k]$, and $\mathbf{E}_{d,j}[k]$ as $\mathbf{W}_{r,n_r} \in \mathbb{C}^{KM \times KM} \succeq \mathbf{0}$, $\mathbf{W}_{u,i}[k] \in \mathbb{C}^{D_i^u \times D_i^u} \succeq \mathbf{0}$, and $\mathbf{W}_{d,j}[k] \in \mathbb{C}^{D_j^d \times D_j^d} \succeq \mathbf{0}$, respectively. The weighted-sum MSE is

$$\begin{aligned} \Xi_{\text{wmse}} &\triangleq \underbrace{\sum_{k=1}^K \sum_{i=1}^I \alpha_i^u \text{Tr} \{ \mathbf{W}_{u,i}[k] \mathbf{E}_{u,i}[k] \}}_{=\Xi_{\text{UL}}} + \underbrace{\sum_{n_r=1}^{N_r} \alpha_{n_r}^r \text{Tr} \{ \mathbf{W}_{r,n_r} \mathbf{E}_{r,n_r} \}}_{=\Xi_r} \\ &\quad + \underbrace{\sum_{k=1}^K \sum_{j=1}^J \alpha_j^d \text{Tr} \{ \mathbf{W}_{d,j}[k] \mathbf{E}_{d,j}[k] \}}_{=\Xi_{\text{DL}}}, \end{aligned} \quad (38)$$

where $\mathbf{W}_{u,i}[k] \mathbf{E}_{u,i}[k]$, $\mathbf{W}_{d,j}[k] \mathbf{E}_{d,j}[k]$ and $\mathbf{W}_{r,n_r} \mathbf{E}_{r,n_r}$ are:

$$\begin{aligned} &\mathbf{W}_{u,i}[k] \mathbf{E}_{u,i}[k] \\ &= \mathbf{W}_{u,i}[k] \mathbf{I} - \mathbf{W}_{u,i}[k] \mathbf{U}_{u,i}[k] \mathbf{H}_{i,B} \mathbf{P}_{u,i}[k] - \mathbf{W}_{u,i}[k] \mathbf{P}_{u,i}^\dagger[k] \mathbf{H}_{i,B}^\dagger \mathbf{U}_{u,i}^\dagger[k] + \mathbf{W}_{u,i}[k] \mathbf{U}_{u,i}[k] \mathbf{H}_{i,B} \mathbf{P}_{u,i}[k] \mathbf{P}_{u,i}^\dagger[k] \mathbf{H}_{i,B}^\dagger \mathbf{U}_{u,i}^\dagger[k] \\ &\quad + \mathbf{W}_{u,i}[k] \mathbf{U}_{u,i}[k] \left(\sum_{q \neq i} \mathbf{H}_{q,B} \mathbf{P}_{u,q}[k] \mathbf{P}_{u,q}^\dagger[k] \mathbf{H}_{q,B}^\dagger + \sum_{j=1}^J \mathbf{H}_{B,j} \mathbf{P}_{d,j}[k] \mathbf{P}_{d,j}^\dagger[k] \mathbf{H}_{B,j}^\dagger + \mathbf{H}_{rB} \mathbf{a}[k] \mathbf{a}^\dagger[k] \mathbf{H}_{rB}^\dagger \right) \mathbf{U}_{u,i}^\dagger[k], \end{aligned} \quad (39)$$

$$\begin{aligned}
 & \mathbf{W}_{d,j}[k] \mathbf{E}_{d,j}[k] \\
 &= \mathbf{W}_{d,j}[k] \mathbf{I} - \mathbf{W}_{d,j}[k] \mathbf{U}_{d,j}[k] \mathbf{H}_{B,j} \mathbf{P}_{d,j}[k] - \mathbf{W}_{d,j}[k] \mathbf{P}_{d,j}^\dagger[k] \mathbf{H}_{B,j}^\dagger \mathbf{U}_{d,j}^\dagger[k] \\
 &+ \mathbf{W}_{d,j}[k] \mathbf{U}_{d,j}[k] \mathbf{H}_{B,j} \mathbf{P}_{d,j}[k] \mathbf{P}_{d,j}^\dagger[k] \mathbf{H}_{B,j}^\dagger \mathbf{U}_{d,j}^\dagger[k] \\
 &+ \mathbf{W}_{d,j}[k] \mathbf{U}_{d,j}[k] \left(\sum_{g \neq j} \mathbf{H}_{B,j} \mathbf{P}_{u,g}[k] \mathbf{P}_{u,g}^\dagger[k] \mathbf{H}_{B,j}^\dagger + \sum_{i=1}^I \mathbf{H}_{i,j} \mathbf{P}_{u,i}[k] \mathbf{P}_{u,i}^\dagger[k] \mathbf{H}_{i,j}^\dagger + \mathbf{H}_{r,j} \mathbf{a}[k] \mathbf{a}^\dagger[k] \mathbf{H}_{r,j}^\dagger \right) \mathbf{U}_{d,j}^\dagger[k], \quad (40)
 \end{aligned}$$

and

$$\begin{aligned}
 & \mathbf{W}_{r,n_r} \mathbf{E}_{r,n_r} \\
 &= \mathbf{W}_{r,n_r} \mathbb{E} \left[\left(\mathbf{h}_{t,n_r} - \sum_{k=1}^K \mathbf{u}_{r,n_r}[k] \left\{ \mathbf{y}_{t,n_r}[k] + \mathbf{y}_{r,n_r}^{\text{in}}[k] \right\} \right) \left(\mathbf{h}_{t,n_r} - \sum_{k=1}^K \mathbf{u}_{r,n_r}[k] \left\{ \mathbf{y}_{t,n_r}[k] + \mathbf{y}_{r,n_r}^{\text{in}}[k] \right\} \right)^\dagger \right] \\
 &= \mathbf{W}_{r,n_r} \boldsymbol{\Sigma}_{t,n_r} - \mathbf{W}_{r,n_r} \mathbb{E} \left[\sum_{m=1}^K \mathbf{J}_h[m] \left(\mathbf{J}_r \mathbf{h}_{rt,n_r}[m] + \mathbf{J}_B \mathbf{h}_{Bt,n_r}[m] \right) \sum_{\ell=1}^K \left(\mathbf{a}^\dagger[\ell] \mathbf{h}_{rt,n_r}^*[\ell] + \mathbf{s}_{Bt,n_r}^\dagger[\ell] \mathbf{h}_{Bt,n_r}^*[\ell] \right) \mathbf{u}_{r,n_r}^\dagger[\ell] \right] \\
 &- \mathbf{W}_{r,n_r} \mathbb{E} \left[\sum_{m=1}^K \mathbf{u}_{r,n_r}[m] \left(\mathbf{h}_{rt,n_r}[m] \mathbf{a}[m] + \mathbf{h}_{Bt,n_r}[m] \mathbf{s}_{Bt,n_r}[m] \right) \sum_{\ell=1}^K \left(\mathbf{h}_{rt,n_r}^\dagger[\ell] \mathbf{J}_r^\dagger + \mathbf{h}_{Bt,n_r}^\dagger[\ell] \mathbf{J}_B^\dagger \right) \mathbf{J}_h^\dagger[\ell] \right] \\
 &+ \mathbf{W}_{r,n_r} \sum_{m=1}^K \sum_{\ell=1}^K \mathbf{u}_{r,n_r}[m] \left(\mathbf{R}_{rt,n_r}(m, \ell) + \mathbf{R}_{Bt,n_r}(m, \ell) + \mathbf{R}_{Bm,n_r}(m, \ell) + \mathbf{R}_{U,n_r}(m, \ell) + \mathbf{R}_{c,n_r}(m, \ell) + \sigma_{n_r} \right) \mathbf{u}_{r,n_r}^\dagger[\ell]. \quad (41)
 \end{aligned}$$

Minimizing (38) is the key to solving the problematic non-convex problem in (34), as stated in the following theorem.

Theorem 1. *Solving the problem*

$$\begin{aligned}
 & \underset{\{\mathbf{P}\}, \{\mathbf{U}\}, \mathbf{A}, \{\mathbf{W}\}}{\text{minimize}} \quad \Xi_{\text{wmse}}(\{\mathbf{U}\}, \{\mathbf{P}\}, \mathbf{A}, \{\mathbf{W}\}), \quad (42) \\
 & \text{subject to} \quad (33\text{b}) - (33\text{e})
 \end{aligned}$$

yields the exact solution of the problem (34).

Proof. The optimization of (42) w.r.t. $\{\mathbf{U}\}$ yields [28]

$$\begin{aligned}
 \mathbf{U}_{r,n_r}^\star &= \arg \min_{\mathbf{U}_{r,n_r}, \forall n_r} \text{Tr} \left\{ \mathbf{W}_{r,n_r} \mathbf{E}_{r,n_r} \right\} \\
 &= \boldsymbol{\Sigma}_{t,n_r} \mathbf{S}_{t,n_r}^\dagger \left(\mathbf{S}_{t,n_r} \boldsymbol{\Sigma}_{t,n_r} \mathbf{S}_{t,n_r}^\dagger + \mathbf{R}_{r,n_r}^{\text{in}} \right)^{-1}, \quad (43)
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{U}_{u,i}^\star[k] &= \arg \min_{\mathbf{U}_{u,i}[k], \forall i,k,l} \text{Tr} \left\{ \mathbf{W}_{u,i}[k] \mathbf{E}_{u,i}[k] \right\} \\
 &= \mathbf{P}_{u,i}^\dagger[k] \mathbf{H}_{i,B}^\dagger (\mathbf{R}_{u,i}[k])^{-1}, \quad (44)
 \end{aligned}$$

$$\begin{aligned}
 \text{and } \mathbf{U}_{d,j}^\star[k] &= \arg \min_{\mathbf{U}_{d,j}[k], \forall j,k,l} \text{Tr} \left\{ \mathbf{W}_{d,j}[k] \mathbf{E}_{d,j}[k] \right\} \\
 &= \mathbf{P}_{d,j}^\dagger[k] \mathbf{H}_{B,j}^\dagger (\mathbf{R}_{d,j}[k])^{-1}. \quad (45)
 \end{aligned}$$

Substituting $\mathbf{U}_{r,n_r}^\star$, $\mathbf{U}_{u,i}^\star[k]$, and $\mathbf{U}_{d,j}^\star[k]$ into MIs in (27)-(29) and MSEs in (35)-(37) yields, respectively, the achievable rates of n_r -th radar Rx, i -th UL UE, and j -th DL UE as

$$R_{r,n_r} = \log \left| \mathbf{I} + \mathbf{S}_{t,n_r} \boldsymbol{\Sigma}_{t,n_r} \mathbf{S}_{t,n_r}^\dagger \mathbf{R}_{\text{in},n_r}^{-1} \right|, \quad (46)$$

$$R_{u,i}[k] = \log \left| \mathbf{I} + \mathbf{R}_{i,B}[k] \left(\mathbf{R}_{u,i}^{\text{in}}[k] \right)^{-1} \right|, \quad (47)$$

and

$$R_{d,j}[k] = \log \left| \mathbf{I} + \mathbf{R}_{B,j}[k] \left(\mathbf{R}_{d,j}^{\text{in}}[k] \right)^{-1} \right|. \quad (48)$$

The corresponding minimum mean-squared-error (MMSE) of n_r -th radar Rx, i -th UL UE, and j -th DL UE are, respectively,

$$\mathbf{E}_{r,n_r}^* = \mathbf{\Sigma}_{t,n_r} \left[\mathbf{I} - \mathbf{S}_{t,n_r}^\dagger \left(\mathbf{R}_{r,n_r} \right)^{-1} \mathbf{S}_{t,n_r} \mathbf{\Sigma}_{t,n_r} \right], \quad (49)$$

$$\mathbf{E}_{u,i}^*[k] = \mathbf{I} - \mathbf{P}_{u,i}^\dagger[k] \mathbf{H}_{i,B}^\dagger \left(\mathbf{R}_{u,i}[k] \right)^{-1} \mathbf{H}_{i,B} \mathbf{P}_{u,i}[k], \quad (50)$$

and

$$\mathbf{E}_{d,j}^*[k] = \mathbf{I} - \mathbf{P}_{d,j}^\dagger[k] \mathbf{H}_{B,j}^\dagger \left(\mathbf{R}_{d,j}[k] \right)^{-1} \mathbf{H}_{B,j} \mathbf{P}_{d,j}[k]. \quad (51)$$

The *data processing inequality* [4, p.34] implies that R_{r,n_r} , $R_{u,i}[k, l]$, and $R_{d,j}[k, l]$ are the upper bounds of I_{r,n_r} , $I_i^u[k, l]$, and $I_j^d[k, l]$, for all n_r , i , and j . It follows that

$$\{\mathbf{U}^*\} \triangleq \left\{ \mathbf{U}_{r,n_r}^*, \mathbf{U}_{u,i}^*[k], \mathbf{U}_{d,j}^*[k], \forall \{n_r, i, j\} \right\}, \quad (52)$$

is also the optimal solution of (34) and, in turn, the original problem (33), whose additional constraints do not affect the solution for $\{\mathbf{U}\}$. Using Woodbury matrix identity [27], the achievable rates are

$$R_{r,n_r} = \log \left| \mathbf{\Sigma}_{t,n_r} \left(\mathbf{E}_{r,n_r}^* \right)^{-1} \right|, \quad R_{u,i}[k] = \log \left| \left(\mathbf{E}_{u,i}^*[k] \right)^{-1} \right|, \\ \text{and } R_{d,j}[k] = \log \left| \left(\mathbf{E}_{d,j}^*[k] \right)^{-1} \right|. \quad (53)$$

Applying the first order optimal condition [38, 27] w.r.t. $\{\mathbf{W}\}$ produces the optimal weight matrices $\{\mathbf{W}^*\}$ as

$$\mathbf{W}_{u,i}^*[k] = \left(\mathbf{E}_{u,i}^*[k] \right)^{-1}, \\ \mathbf{W}_{d,j}^*[k] = \left(\mathbf{E}_{d,j}^*[k] \right)^{-1}, \quad (54)$$

and

$$\mathbf{W}_{r,n_r}^* = \left(\mathbf{E}_{r,n_r}^* \right)^{-1}, \quad (55)$$

for all $\{n_r, i, j\}$. Define

$$\Xi'_{\text{wmse}} = \Xi_{\text{wmse}} - \sum_{n_r=1}^{N_r} \alpha_{n_r}^r \left(\mathbf{\Sigma}_{t,n_r} \log \left| \mathbf{W}_{r,n_r} \right| + KM \right) - \\ \sum_{k=0}^{K-1} \left\{ \sum_{i=1}^I \alpha_i^u \left(\log \left| \mathbf{W}_{u,i}[k] \right| + D_i^u \right) + \sum_{j=1}^J \alpha_j^d \left(\log \left| \mathbf{W}_{d,j}[k] \right| + D_j^d \right) \right\}.$$

Substituting $\{\mathbf{W}^*\}$ and $\{\mathbf{U}^*\}$ in Ξ'_{wmse} results in

$$\Xi'_{\text{wmse}} = \sum_{k=0}^{K-1} \left\{ - \sum_{j=1}^J \alpha_j^d \log \left| \left(\mathbf{E}_{d,j}^*[k] \right)^{-1} \right| \right\}$$

$$\begin{aligned}
 & - \sum_{i=1}^I \alpha_i^u \log \left| \left(\mathbf{E}_{u,i}^*[k] \right)^{-1} \right| \Bigg\} - \alpha_r \sum_{n_r=1}^{N_r} \log \left| \left(\mathbf{E}_{r,n_r}^* \right)^{-1} \right| \\
 & = \sum_{k=1}^K \left\{ \sum_{j=1}^J \alpha_j^d R_{d,j}[k] + \sum_{i=1}^I \alpha_i^u R_{u,i}[k] \right\} + \sum_{n_r=1}^{N_r} \alpha_{n_r}^r R_{r,n_r} \\
 & = -I_{\text{CWSM}}(\{\mathbf{U}^*\}, \{\mathbf{P}\}, \mathbf{A}),
 \end{aligned} \tag{56}$$

which indicates that maximizing I_{CWSM} is equivalent to minimizing Ξ'_{wmmse} given $\{\mathbf{U}^*\}$. As minimizing Ξ'_{wmmse} w.r.t. $\{\mathbf{P}\}$ and $\mathbf{A}$ is also equivalent to minimizing Ξ_{wmmse} given $\{\mathbf{U}^*\}$ and $\{\mathbf{W}^*\}$. This completes the proof. $\square$

Substituting $\{\mathbf{U}^*\}$ and $\{\mathbf{W}^*\}$ in (42) yields the WMMSE

$$\Xi_{\text{wmmse}}(\{\mathbf{P}\}, \mathbf{A}) \triangleq \Xi_{\text{wmmse}}(\{\mathbf{U}^*\}, \{\mathbf{W}^*\}|\{\mathbf{P}\}). \tag{57}$$

Therefore, $\mathbf{A}$ and $\{\mathbf{P}\}$ are obtained by solving

$$\underset{\{\mathbf{P}\}, \mathbf{A}}{\text{minimize}} \quad \Xi_{\text{wmmse}}(\{\mathbf{P}\}, \mathbf{A}) \text{ subject to } (33\text{b}) - (33\text{e}). \tag{58}$$

4.2. WMMSE-MRMC

In order to solve (58), we sequentially iterate over each element in $\{\mathbf{P}\}$ and each row of $\mathbf{A}$, i.e., $\mathbf{a}^\top[k]$, using the Lagrange dual method to find a closed-form solution to each variable, which constitutes our WMMSE-MRMC algorithm.

4.2.1. Lagrange dual solution

Denote the Lagrange multiplier vectors w.r.t. constraints (33b)-(33e), respectively, as

$$\lambda_{\text{DL}} \triangleq [\lambda_d[0], \dots, \lambda_d[K-1]]^\top \in \mathbb{R}^K, \tag{59}$$

$$\lambda_{\text{UL}} \triangleq [\lambda_{u,1}[0], \dots, \lambda_{u,I}[K-1]]^\top \in \mathbb{R}^{KI}, \tag{60}$$

$$\mu_{\text{DL}} \triangleq [\mu_{d,1}[0], \dots, \mu_{d,J}[K-1]]^\top \in \mathbb{R}^{KJ}, \tag{61}$$

and

$$\mu_{\text{UL}} \triangleq [\mu_{u,1}[0], \dots, \mu_{u,I}[K-1]]^\top \in \mathbb{R}^{KI}, \tag{62}$$

as well as the UL power vector, the DL power vector, the UL rate vector, and the DL rate vector as, respectively,

$$\mathbf{p}_{\text{UL}} \triangleq [P_{u,1}[0], \dots, P_{u,I}[K-1]]^\top \in \mathbb{R}^{KI}, \tag{63}$$

$$\mathbf{p}_{\text{DL}} \triangleq [P_d[0], \dots, P_d[K-1]]^\top \in \mathbb{R}^K, \tag{64}$$

$$\mathbf{r}_{\text{UL}} \triangleq [R_{u,1}[0], \dots, R_{u,I}[K-1]]^\top \in \mathbb{R}^{KI}, \tag{65}$$

and

$$\mathbf{r}_{\text{DL}} \triangleq [R_{d,1}[0], \dots, R_{d,J}[K-1]]^\top \in \mathbb{R}^{KJ}, \tag{66}$$

which lead to the Lagrangian associated with (58) as

$$\begin{aligned}
 \mathcal{L}(\{\mathbf{P}\}, \mathbf{A}, \lambda, \mu) &= \Xi_{\text{wmmse}} + \lambda_{\text{DL}}^\top (\mathbf{p}_{\text{DL}} - P_B \mathbf{1}) + \lambda_{\text{UL}}^\top (\mathbf{p}_{\text{UL}} - P_U \mathbf{1}) \\
 &\quad - \mu_{\text{DL}}^\top (\mathbf{r}_{\text{DL}} - R_{\text{DL}} \mathbf{1}) - \mu_{\text{UL}}^\top (\mathbf{r}_{\text{UL}} - R_{\text{UL}} \mathbf{1}),
 \end{aligned} \tag{67}$$

where $\lambda = [\lambda_{\text{DL}}^\top, \lambda_{\text{UL}}^\top]^\top$ and $\mu = [\mu_{\text{DL}}^\top, \mu_{\text{UL}}^\top]^\top$. The Lagrange dual function of $L(\cdot)$ is defined as $D(\lambda, \mu) = \inf_{\{\mathbf{P}\}, \mathbf{A}} \mathcal{L}(\{\mathbf{P}\}, \mathbf{A}, \lambda, \mu)$. With these definitions, we state the following theorem to solve (58).

Proposition 1. *Linearize, using Taylor series approximation, $R_{u,i}[k]$ and $R_{d,j}[k]$, $\forall \{i, j, k\}$ in the QoS constraints of problem (58). Then, the solution of the resulting problem is equivalent to that of its Lagrange dual problem*

$$\text{maximize } D(\lambda, \mu) \text{ subject to } \lambda \geq \mathbf{0}, \mu \geq \mathbf{0}. \quad (68)$$

Proof. Solving (68) yields the lower bounds of (58). The difference between the lower bound and the actual optimal value is the optimal duality gap. To equivalently obtain $\{\mathbf{P}^*\}$ and $\mathbf{A}'$ with (68), strong duality ought to hold for the primal problem (42), i.e., the optimal duality gap should be zero. However, the QoS constraints (33e) are non-concave leading to a non-zero duality gap. To bypass this problem, we apply the Taylor series to obtain linear approximations of $R_{u,i}[k]$ and $R_{d,j}[k]$. The first-order Taylor series expansion of a real-valued function with complex-valued matrix arguments $f(\mathbf{X}, \mathbf{X}^*) : \mathbb{C}^{N \times Q} \times \mathbb{C}^{N \times Q} \rightarrow \mathbb{R}$ around $\mathbf{X}_0$ yields [27]

$$\begin{aligned} f(\mathbf{X}, \mathbf{X}^*) &= f(\mathbf{X}_0, \mathbf{X}_0^*) + \text{vec}^\top \left(\frac{\partial}{\partial \mathbf{X}_0} f(\mathbf{X}_0, \mathbf{X}_0^*) \right) \text{vec}(\mathbf{X} - \mathbf{X}_0) \\ &+ \text{vec}^\top \left(\frac{\partial}{\partial \mathbf{X}_0^*} f(\mathbf{X}_0, \mathbf{X}_0^*) \right) \text{vec}(\mathbf{X}^* - \mathbf{X}_0^*). \end{aligned} \quad (69)$$

based on their associated Taylor series expansions in the initial approximations $\tilde{\mathbf{P}}_{u,i}[k]$, $\tilde{\mathbf{P}}_{d,j}[k]$, and $\tilde{\mathbf{a}}[k]$. Denoting $\tilde{\mathbf{P}}_{u,i}[k]$ as an initial approximation of $\mathbf{P}_{u,i}[k]$, the Taylor series expansions of $R_{u,q}[k]$ at $\tilde{\mathbf{P}}_{u,i}[k]$ are

$$\begin{aligned} R_{u,q}[k](\mathbf{P}_{u,i}[k]) &\approx R_{u,q}[k](\tilde{\mathbf{P}}_{u,i}[k]) \\ &+ \text{vec} \left\{ \frac{\partial \left(R_{u,q}[k](\tilde{\mathbf{P}}_{u,i}[k], \tilde{\mathbf{P}}_{u,i}^*[k]) \right)}{\partial \tilde{\mathbf{P}}_{u,i}[k]} \right\}^\top \text{vec} \left\{ \mathbf{P}_{u,i}[k] - \tilde{\mathbf{P}}_{u,i}[k] \right\} \\ &+ \text{vec} \left\{ \frac{\partial \left(R_{u,q}[k](\tilde{\mathbf{P}}_{u,i}[k], \tilde{\mathbf{P}}_{u,i}^*[k]) \right)}{\partial \tilde{\mathbf{P}}_{u,i}^*[k]} \right\}^\top \text{vec} \left\{ \mathbf{P}_{u,i}^*[k] - \tilde{\mathbf{P}}_{u,i}^*[k] \right\}. \end{aligned}$$

Likewise, by defining $\tilde{\mathbf{P}}_{d,j}[k]$ and $\tilde{\mathbf{a}}[k]$ as initial approximations of $\mathbf{P}_{d,j}[k]$, and $\mathbf{a}[k]$, we find the linear approximations of $R_{u,q}[k](\mathbf{P}_{d,j}[k])$, $R_{u,q}[k](\mathbf{a}[k])$, $R_{d,j}[k](\mathbf{P}_{u,i}[k])$, $R_{d,m}[k](\mathbf{P}_{d,j}[k])$, and $R_{d,j}[k](\mathbf{a}[k])$. As Ξ_{wmmse} is multi-convex, namely that Ξ_{wmmse} is not jointly convex in $\mathbf{P}_{u,i}[k]$, $\mathbf{P}_{d,j}[k]$, and $\mathbf{a}[k]$ but convex in each individual variable provided the rest of the variables are fixed [28, 31], we thus have a fully convex approximation of (58) in (68) and reduce the optimal duality gap to zero [30]. This concludes the proof. $\square$

Following Appendix A, the gradients of Lagrangian $\mathcal{L}(\cdot)$ w.r.t. $\mathbf{P}_{u,i}[k]$, $\mathbf{P}_{d,j}[k]$, and $\mathbf{a}[k]$ are, respectively,

$$\begin{aligned} \nabla_{\mathbf{P}_{u,i}[k]} \mathcal{L} &= \nabla_{\mathbf{P}_{u,i}[k]} \Xi_{\text{UL}} + \nabla_{\mathbf{P}_{u,i}[k]} \Xi_{\text{DL}} + \nabla_{\mathbf{P}_{u,i}[k]} \Xi_{\text{r}} \\ &+ \lambda_{u,i}[k] \mathbf{P}_{u,i}[k] - \sum_{g=1}^J \mu_{d,g}[k] \nabla_{\mathbf{P}_{u,i}[k]} R_{d,g}[k] - \sum_{q=1}^I \mu_{u,q}[k] \nabla_{\mathbf{P}_{u,i}[k]} R_{u,q}[k], \\ \nabla_{\mathbf{P}_{d,j}[k]} \mathcal{L} &= \nabla_{\mathbf{P}_{d,j}[k]} \Xi_{\text{UL}} + \nabla_{\mathbf{P}_{d,j}[k]} \Xi_{\text{DL}} + \nabla_{\mathbf{P}_{d,j}[k]} \Xi_{\text{r}} \\ &+ \lambda_{d,j}[k] \mathbf{P}_{d,j}[k] - \sum_{g=1}^J \mu_{d,g}[k] \nabla_{\mathbf{P}_{d,j}[k]} R_{d,g}[k] - \sum_{q=1}^I \mu_{u,q}[k] \nabla_{\mathbf{P}_{d,j}[k]} R_{u,q}[k], \\ \text{and } \nabla_{\mathbf{a}[k]} \mathcal{L} &= \nabla_{\mathbf{a}[k]} \Xi_{\text{UL}} + \nabla_{\mathbf{a}[k]} \Xi_{\text{DL}} + \nabla_{\mathbf{a}[k]} \Xi_{\text{r}} - \sum_{j=1}^J \mu_{d,j}[k] \nabla_{\mathbf{a}[k]} R_{d,j}[k] - \sum_{i=1}^I \mu_{u,i}[k] \nabla_{\mathbf{a}[k]} R_{u,i}[k]. \end{aligned}$$

We obtain the analytical solutions to $\mathbf{P}_{u,i}[k]$, $\mathbf{P}_{d,j}[k]$, and $\mathbf{a}[k]$ by solving equations $\nabla_{\mathbf{P}_{u,i}[k]}L = 0$, $\nabla_{\mathbf{P}_{d,j}[k]}L = 0$, and $\nabla_{\mathbf{a}[k]}L = 0$, respectively. Using the gradient expressions in Appendix A, some algebra yields three generalized Sylvester equations

$$\mathbf{A}_{u,i}[k]\mathbf{P}_{u,i}[k] + \mathbf{F}_{u,i}[k]\mathbf{P}_{u,i}[k]\mathbf{B}_{u,i}[k] = \mathbf{C}_{u,i}[k], \quad (70)$$

$$\mathbf{A}_{d,j}[k]\mathbf{P}_{d,j}[k] + \mathbf{F}_{Bt}[k]\mathbf{P}_{d,j}[k]\mathbf{B}_{Bt,j}[k] + \mathbf{F}_{Bm}[k]\mathbf{P}_{d,j}[k]\mathbf{B}_{Bm,j}[k] = \mathbf{C}_{d,j}[k], \quad (71)$$

$$\text{and } \mathbf{A}_r[k]\mathbf{a}[k] + \mathbf{F}_r[k]\mathbf{a}[k] = \mathbf{c}_r[k], \quad (72)$$

where

$$\mathbf{B}_{u,i}[k] = \mathbf{d}_{u,i}[k, n_t - n_u]\mathbf{d}_{u,i}^\dagger[k, n_t - n_u],$$

$$\mathbf{F}_{u,i}[k] = 2 \sum_{n_r=1}^{N_r} \text{Re} \left(\xi_{r,n_r}(k, k) \mathbf{\Sigma}_{i,n_r}^{(k,k)} \right),$$

$$\mathbf{B}_{Bt,j}[k] = \mathbf{d}_{d,j}[k, 0]\mathbf{d}_{d,j}^\dagger[k, 0],$$

$$\mathbf{B}_{Bm,j}[k] = \mathbf{d}_{d,j}[k, n_t - n_{Bm}]\mathbf{d}_{d,j}^\dagger[k, n_t - n_{Bm}],$$

$$\mathbf{F}_{Bt}[k] = 2 \sum_{n_r=1}^{N_r} \text{Re} \left(\xi_{r,n_r}(k, k) \mathbf{\Sigma}_{Bt,n_r}^{(k,k)} \right),$$

$$\mathbf{F}_{Bm,j}[k] = 2 \sum_{n_r=1}^{N_r} \text{Re} \left(\xi_{r,n_r}(k, k) \mathbf{\Sigma}_{Bm,n_r}^{(k,k)} \right),$$

$$\mathbf{F}_r[k] = 2 \sum_{n_r=1}^{N_r} \left[\text{Re} \left(\xi_{r,n_r}(k, k) \mathbf{\Sigma}_{rt,n_r}^{(k,k)} \right) + \text{Re} \left(\xi_{r,n_r}(k, k) \mathbf{\Sigma}_{c,n_r} \right) \right],$$

$$\mathbf{F}_r[k] = 2 \sum_{n_r=1}^{N_r} \left[\text{Re} \left(\xi_{r,n_r}(k, k) \mathbf{\Sigma}_{rt,n_r}^{(k,k)} \right) + \text{Re} \left(\xi_{r,n_r}(k, k) \mathbf{\Sigma}_{c,n_r} \right) \right],$$

$$\mathbf{A}_{u,i}[k] = 2\mathbf{H}_{i,B}^\dagger \xi_{UL}[k]\mathbf{H}_{i,B} + 2 \sum_{g=1}^J \mathbf{H}_{i,g}^\dagger \xi_{d,g}[k]\mathbf{H}_{i,g} + \lambda_{u,i}[k]\mathbf{I},$$

$$\begin{aligned} \mathbf{C}_{u,i}[k] &= \sum_{q=1}^I \mu_{u,q}[k] \nabla_{\mathbf{P}_{u,i}[k]} R_{u,q}[k] + \sum_{g=1}^J \mu_{d,g}[k] \nabla_{\mathbf{P}_{u,i}[k]} R_{d,g}[k] \\ &\quad - \sum_{n_r=1}^{N_r} \sum_{m \neq k}^K 2\text{Re} \left(\xi_{r,n_r}(k, m) \mathbf{\Sigma}_{i,n_r}^{(m,k)} \right) \mathbf{P}_{u,i}[m] \mathbf{d}_{u,i}[m, n_t - n_u] \mathbf{d}_{u,i}^\dagger[k, n_t - n_u] + 2\alpha_i^u \mathbf{H}_{i,B}^\dagger \mathbf{U}_{u,i}^\dagger[k] \mathbf{W}_{u,i}[k], \end{aligned}$$

$$\mathbf{A}_{d,j}[k] = 2\mathbf{H}_{BB}^\dagger \xi_{UL}[k]\mathbf{H}_{BB} + 2 \sum_{g=1}^J \mathbf{H}_{B,g}^\dagger \xi_{d,g}[k]\mathbf{H}_{B,g} + \lambda_d[k]\mathbf{I},$$

$$\begin{aligned} \mathbf{C}_{d,j}[k] &= \sum_{q=1}^I \mu_{u,q}[k] \nabla_{\mathbf{P}_{d,j}[k]} R_{u,q}[k] + \sum_{g=1}^J \mu_{d,g}[k] \nabla_{\mathbf{P}_{d,j}[k]} R_{d,g}[k] - \\ &\quad \sum_{n_r=1}^{N_r} \left\{ \sum_{m \neq k} 2\text{Re} \left(\xi_{r,n_r}(m, k) \mathbf{\Sigma}_{Bt,n_r}^{(k,m)} \right) \sum_{g=1}^J \mathbf{P}_{d,g}[m] \mathbf{d}_{d,g}[m, 0] \mathbf{d}_{d,j}^\dagger[k, 0] \right. \\ &\quad \left. + \sum_{m \neq k} 2\text{Re} \left(\xi_{r,n_r}(m, k) \mathbf{\Sigma}_{Bm,n_r}^{(k,m)} \right) \sum_{g=1}^J \mathbf{P}_{d,g}[m] \mathbf{d}_{d,g}[m, n_t - n_{Bm}] \mathbf{d}_{d,j}^\dagger[k, n_t - n_{Bm}] \right\} \end{aligned}$$

$$\begin{aligned}
 & + \sum_{g \neq j} 2\text{Re}\left(\xi_{r,n_r}(k, k) \Sigma_{\text{Bt},n_r}^{(k,k)}\right) \mathbf{P}_{\text{d},g}[k] \mathbf{d}_{\text{d},g}[k, 0] \mathbf{d}_{\text{d},j}^\dagger[k, 0] \\
 & + \sum_{g \neq j} 2\text{Re}\left(\xi_{r,n_r}(k, k) \Sigma_{\text{Bm},n_r}^{(k,k)}\right) \mathbf{P}_{\text{d},g}[k] \mathbf{d}_{\text{d},g}[k, n_t - n_{\text{Bm}}] \mathbf{d}_{\text{d},j}^\dagger[k, n_t - n_{\text{Bm}}] \\
 & - 2\alpha_{n_r}^r \sum_{m=1}^K \Sigma_{\text{Bt},n_r}^{(m,k)} \mathbf{J}_{\text{B}}^\top \mathbf{J}_{\text{h}}^\top[m] \mathbf{W}_{r,n_r} \mathbf{u}_{r,n_r}[k] \mathbf{d}_{\text{d},j}^\dagger[k, 0] \Big\} + 2\alpha_j^d \mathbf{H}_{\text{B},j}^\dagger \mathbf{U}_{\text{d},j}^\dagger[k] \mathbf{W}_{\text{d},j}[k], \\
 \mathbf{A}_r[k] & = 2\mathbf{H}_{\text{rB}}^\dagger \xi_{\text{UL}}[k] \mathbf{H}_{\text{rB}} + 2 \sum_{j=1}^J \mathbf{H}_{r,j}^\dagger \xi_{\text{d},j}[k] \mathbf{H}_{r,j}, \\
 \text{and } \mathbf{c}_r[k] & = \sum_{i=1}^I \mu_{u,i}[k] \nabla_{\mathbf{a}[k]} R_{u,i}[k] + \sum_{j=1}^J \mu_{d,j}[k] \nabla_{\mathbf{a}[k]} R_{d,j}[k] + 2 \sum_{n_r=1}^{N_r} \alpha_{r,n_r} \sum_{m=1}^K \text{Re}\left(\Sigma_{\text{rt},n_r}^{(m,k)}\right) \mathbf{J}_{\text{r}}^\top \mathbf{J}_{\text{h}}^\top[m] \mathbf{W}_{r,n_r} \mathbf{u}_{r,n_r}[k] \\
 & - 2 \sum_{n_r=1}^{N_r} \sum_{m \neq k} \left[\text{Re}\left(\xi_{r,n_r}(k, m) \Sigma_{\text{rt},n_r}^{(m,k)}\right) + \text{Re}\left(\xi_{r,n_r}(k, m) \Sigma_{\text{c},n_r}\right) \right] \mathbf{a}[m].
 \end{aligned}$$

Solving these Sylvester equations yields [27]

$$\text{vec}\left(\mathbf{P}_{u,i}^*[k]\right) = \left[\mathbf{I}_{D_i^u} \otimes \mathbf{A}_{u,i}[k] + \mathbf{B}_{u,i}^\top[k] \otimes \mathbf{F}_{u,i}[k]\right]^{-1} \text{vec}\left(\mathbf{C}_{u,i}[k]\right), \quad (73a)$$

$$\text{vec}\left(\mathbf{P}_{d,j}^*[k]\right) = \left(\mathbf{I} \otimes \mathbf{A}_{d,j}[k] + \mathbf{B}_{\text{Bt},j}^\top[k] \otimes \mathbf{F}_{\text{Bt},j}[k] + \mathbf{B}_{\text{Bm},j}^\top[k] \otimes \mathbf{F}_{\text{Bm},j}[k]\right)^{-1} \text{vec}\left(\mathbf{C}_{d,j}[k]\right), \quad (73b)$$

$$\text{and } \mathbf{a}'[k] = \left[\mathbf{I} \otimes \mathbf{A}_r[k] + \mathbf{I} \otimes \mathbf{F}_r[k]\right]^{-1} \mathbf{c}_r[k], \quad (73c)$$

for all i, j, k . In order to determine $\mathbf{P}_{u,i}^*[k]$, $\mathbf{P}_{d,j}^*[k]$, and $\mathbf{a}'[k]$, we need to find the optimal λ and μ denoted by λ^* and μ^* .

4.2.2. Sub-gradient method for precoders

As $D(\lambda, \mu)$ is not always differentiable [30] and a simple method to update λ and μ is needed, we resort to the subgradient to determine search directions for λ and μ , and employ the projected subgradient method to solve (68) sequentially. In the t -th iteration we have,

$$\lambda_{u,i}^{(t+1)}[k] = \left[\lambda_{u,i}^{(t)}[k] + \beta_{u,i}^{(t)}[k] \left(P_{u,i}^{(t)}[k] - P_U\right)\right]^+, \quad (74a)$$

$$\lambda_d^{(t+1)}[k] = \left[\lambda_d^{(t)}[k] + \beta_d^{(t)}[k] \left(P_B^{(t)}[k] - P_B\right)\right]^+, \quad (74b)$$

$$\mu_{u,i}^{(t+1)}[k] = \left[\mu_{u,i}^{(t)}[k] + \epsilon_{u,i}^{(t)}[k] \left(R_{\text{UL}} - R_{u,i}^{(t)}[k]\right)\right]^+, \quad (74c)$$

$$\text{and } \mu_{d,j}^{(t+1)}[k] = \left[\mu_{d,j}^{(t)}[k] + \epsilon_{d,j}^{(t)}[k] \left(R_{\text{DL}} - R_{d,j}^{(t)}[k]\right)\right]^+, \quad (74d)$$

where $\beta_{u,i}^{(t)}[k]$, $\beta_d^{(t)}[k]$, $\epsilon_{u,i}^{(t)}[k]$, and $\epsilon_{d,j}^{(t)}[k]$ denote the step sizes of the t -th iteration for $\lambda_{u,i}^{(t)}[k]$, $\lambda_d^{(t)}[k]$, $\mu_{u,i}^{(t)}[k]$, and $\mu_{d,j}^{(t)}[k]$, respectively, $P_{u,i}^{(t)}[k]$, $P_B^{(t)}[k]$, $\mathbf{P}_{u,i}^{(t)}[k]$, $\mathbf{P}_{d,j}^{(t)}[k]$, $R_{u,i}^{(t)}[k]$, $R_{d,j}^{(t)}[k]$, and $\Xi_{\text{wmmse}}^{(t)}$ the t -th iterates of $P_{u,i}[k]$, $P_B[k]$, $\mathbf{P}_{u,i}[k]$, $\mathbf{P}_{d,j}[k]$, $R_{u,i}[k]$, $R_{d,j}[k]$, and Ξ_{wmmse} , respectively [30]. Note that $\mathbf{P}_{u,i}^{(t)}[k]$ and $\mathbf{P}_{d,j}^{(t)}[k]$ are obtained by replacing $\lambda_{u,i}[k]$ and $\mu_{u,i}[k]$ with $\lambda_{u,i}^{(t)}[k]$ and $\mu_{u,i}^{(t)}[k]$ in (73a), and $\lambda_d[k]$ and $\mu_{d,j}[k]$ with $\lambda_d^{(t)}[k]$ and $\mu_{d,j}^{(t)}[k]$ in (73b), respectively.

There are various options to choose $\beta_{u,i}^{(t)}[k]$, $\beta_d^{(t)}[k]$, $\epsilon_{u,i}^{(t)}[k]$, and $\epsilon_{d,j}^{(t)}[k]$ to ensure that the subgradient updates in (74a)-(74d) converge to optimal values λ^* and μ^* of λ and μ , respectively. Broadly, two rules are used. The first

determines the step size before executing the algorithm. This includes fixed and progressively diminishing step sizes; the latter should be square summable (but not necessarily summable). The second rule computes the steps using, for example, the Polyak's procedure [39]. As discussed in [39] and [30], the convergence rate of subgradient is dependent on step sizes and initialization points. It was mentioned in [39] that the Polyak's rule achieves faster convergence rate than other rules because it utilizes the optimal function value of the current iteration while computing the step size. Applying the Polyak's rule to $\beta_{u,i}^{(t)}[k]$, $\beta_d^{(t)}[k]$, $\varepsilon_{u,i}^{(t)}[k]$, and $\varepsilon_{d,j}^{(t)}[k]$ produces

$$\beta_{u,i}^{(t)}[k] = (\Xi_{\text{wmmse}}^{(t)} - \Xi_{\text{wmmse}}^{\min} + 0.1^t) / \left| P_{u,i}^{(t)}[k] - P_U \right|^2, \quad (75a)$$

$$\beta_d^{(t)}[k] = (\Xi_{\text{wmmse}}^{(t)} - \Xi_{\text{wmmse}}^{\min} + 0.1^t) / \left| P_B^{(t)}[k] - P_B \right|^2, \quad (75b)$$

$$\varepsilon_{u,i}^{(t)}[k] = (\Xi_{\text{wmmse}}^{(t)} - \Xi_{\text{wmmse}}^{\min} + 0.1^t) / \left| R_{UL} - R_{u,i}^{(t)}[k] \right|^2, \quad (75c)$$

$$\text{and } \varepsilon_{d,j}^{(t)}[k] = (\Xi_{\text{wmmse}}^{(t)} - \Xi_{\text{wmmse}}^{\min} + 0.1^t) / \left| R_{DL} - R_{d,j}^{(t)}[k] \right|^2. \quad (75d)$$

Denote $\ell_{\max}$, $t_{\max}$, $t_{u,\max}$ ($t_{d,\max}$) denote the maximum iterations for the BCD-AP MRMC, WMMSE-MRMC, and the subgradient algorithms, respectively; $(\cdot)^{(\ell,t)}$ as the iterate of a variable at the ℓ -th, t -th, and t -th iterations of BCD-AP MRMC, WMMSE-MRMC, and subgradient algorithms; and $\{\mathbf{P}^{(\ell,t,0)}\} \triangleq \{\mathbf{P}_{u,i}^{(\ell,t,0)}[k], \mathbf{P}_{d,j}^{(\ell,t,0)}[k], \forall \{i, j, k\}\}$. The subgradient method is not descent-based and $\Xi_{\text{wmmse}}^{(t)}$ may increase at certain iterations[30]. Therefore, $\lambda_{u,i}^*[k]$, $\mu_{u,i}^*[k]$, $\lambda_d^*[k]$, and $\mu_{d,j}^*[k]$ are employed to keep track of those values of $\lambda_{u,i}^{(t)}[k]$, $\mu_{u,i}^{(t)}[k]$, $\lambda_d^{(t)}[k]$, and $\mu_{d,j}^{(t)}[k]$ that yield the minimum $\Xi_{\text{wmmse}}^{(t)}$ in the current iteration, i.e. $\Xi_{\text{wmmse}}^{\min}$. Algorithm 1 summarizes the steps of the sub-gradient-based procedures to find $\lambda_{u,i}[k]$, $\mu_{u,i}[k]$ and corresponding $\mathbf{P}_{u,i}^{(\ell,t)}[k]$. We set $\tilde{\mathbf{P}}_{u,i}[k]$ as the optimal estimates of the previous iteration (step 5). We use the same procedure for DL UE to find $\lambda_d[k]$, $\mu_{d,j}[k]$ and $\mathbf{P}_{d,j}^{(\ell,t)}[k]$.

Algorithm 1 Subgradient approach to solve (68) for UL UE

Input: $\{\mathbf{P}^{(\ell,t)}\}$, $\mathbf{A}^{(\ell,t)}$, $\{\mathbf{U}^{(\ell)}\}$, $t_{u,\max}$
Output: $\mathbf{P}_{u,i}^{(\ell,t)}[k]$, $\mu_{u,i}^*[k]$

- 1: Initialize $\lambda_{u,i}^{(0)}[k] = 1$, $\mu_{u,i}^{(0)}[k] = 1$, $\{\mathbf{P}^{(\ell,t,0)}\} = \{\mathbf{P}^{(\ell,t)}\}$
- 2: $t \leftarrow 1$, $\Xi_{\text{wmmse}}^{(\min)} \leftarrow \frac{\{\mathbf{P}^{(\ell,t)}\}, \mathbf{A}^{(\ell,t)}, \{\mathbf{U}^{(\ell)}\}}{(38)}$
- 3: **repeat**
- 4: update $P_{u,i}^{(t)}[k]$, $R_{u,i}^{(t)}[k]$, $\beta_{u,i}^{(t)}[k]$, $\varepsilon_{u,i}^{(t)}[k]$, $\lambda_{u,i}^{(t)}[k]$, $\mu_{u,i}^{(t)}[k]$
- 5: $\tilde{\mathbf{P}}_{u,i}[k] = \mathbf{P}_{u,i}^{(\ell,t)}[k]$ (73a)
- 6: $\Xi_{\text{wmmse}}^{(t)} \leftarrow \frac{\{\mathbf{P}^{(\ell,t,t)}\}, \mathbf{A}^{(\ell,t)}, \{\mathbf{U}^{(\ell)}\}}{(38)}$
- 7: **if** $\Xi_{\text{wmmse}}^{\min} > \Xi_{\text{wmmse}}^{(t)}$ **then** $\mathbf{P}_{u,i}^{(\ell,t)}[k] = \mathbf{P}_{u,i}^{(\ell,t,t)}[k]$, $\lambda_{u,i}^*[k] = \lambda_{u,i}^{(t)}[k]$, $\mu_{u,i}^*[k] = \mu_{u,i}^{(t)}[k]$, and $\Xi_{\text{wmmse}}^{\min} = \Xi_{\text{wmmse}}^{(t)}$
- 8: $t \leftarrow t + 1$
- 9: **until** $t > t_{u,\max}$
- 10: **return** $\mathbf{P}_{u,i}^{(\ell,t)}[k]$, $\mu_{u,i}^*[k]$

Upon executing the subgradient algorithms for all i, j , and k , $\mathbf{A}^{(\ell,t)}$ is solved by replacing μ^* and $\{\mathbf{P}^{(\ell,t)}\}$ with μ and $\{\mathbf{P}\}$ in (73c) during the t -th iteration of the WMMSE-MRMC algorithm (Algorithm 2), whose maximum number of iterations is $t_{\max}$. The outputs of the Algorithm 2 constitute the ℓ -th iterate of $\{\mathbf{P}^{(\ell)}\}$ and $\mathbf{A}'$ for the Algorithm 4.

So far, Algorithm 2 utilized perfect channel state information (CSI). In practice, the CSI is estimated. The resulting estimation error may be modeled either as norm-bounded or stochastically [28]. The former is employed when quantization error is the primary source of the CSI error. However, quantization analysis is beyond the scope of this paper. Therefore, we adopt the latter by modeling the channel matrix as $\hat{\mathbf{H}} = \mathbf{H} + \Delta$, where $\Delta \sim \mathcal{CN}(\mathbf{0}, \eta_{\text{CSI}}^2 \mathbf{I})$ is the error with the variance $\eta_{\text{CSI}}^2 \mathbf{I}$.

Algorithm 2 WMMSE-MRMC algorithm to solve (58)

Input: $\{\mathbf{P}^{(\ell)}\}, \mathbf{A}^{(\ell)}, \{\mathbf{U}^{(\ell)}\}, l_{\max}, t_{u,\max},$ and $t_{d,\max}$
Output: $\{\mathbf{P}^{(\ell)}\}, \mathbf{A}'$

- 1: Set $\{\mathbf{P}^{(\ell,0)}\} \triangleq \{\mathbf{P}_{u,i}^{(\ell,0)}[k], \mathbf{P}_{d,j}^{(\ell,0)}[k], \forall \{i, j, k\}\} = \{\mathbf{P}^{(\ell)}\}$ and $\mathbf{A}^{(\ell,0)} \triangleq \left[(\mathbf{a}^{(\ell,0)}[0])^\top; \dots; (\mathbf{a}^{(\ell,0)}[K])^\top \right] = \mathbf{A}^{(\ell)}$
- 2: Set the iteration index $\iota = 0$
- 3: **repeat**
- 4: **for** $k = 1, \dots, K$ **do**
- 5: **for** $i = 1, \dots, I, j = 1, \dots, J$ **do**
- 6: $\mathbf{P}_{u,i}^{(\ell,\iota+1)}[k], \mu_{u,i}^\star[k] \xleftarrow[t_{u,\max}, \{\mathbf{U}^{(\ell)}\}]{\text{Subgradient}} \{\mathbf{P}^{(\ell,\iota)}\}, \mathbf{A}^{(\ell,\iota)}$
- 7: $\mathbf{P}_{d,j}^{(\ell,\iota+1)}[k], \mu_{d,j}^\star[k] \xleftarrow[t_{d,\max}, \{\mathbf{U}^{(\ell)}\}]{\text{Subgradient}} \{\mathbf{P}^{(\ell,\iota)}\}, \mathbf{A}^{(\ell,\iota)}$
- 8: $\mathbf{a}^{(\ell,\iota+1)}[k] \xleftarrow{(73c)} \{\mathbf{P}^{(\ell,\iota)}\}, \mathbf{A}^{(\ell,\iota)}, \{\mathbf{U}^{(\ell)}\}, \mu^\star$
- 9: $\mathbf{A}^{(\ell,\iota+1)} = \left[(\mathbf{a}^{(\ell,\iota+1)}[0])^\top; \dots; (\mathbf{a}^{(\ell,\iota+1)}[K])^\top \right]$
- 10: $\iota \leftarrow \iota + 1$
- 11: **until** $\iota > l_{\max}$
- 12: $\{\mathbf{P}^{(\ell)}\} \leftarrow \{\mathbf{P}^{(\ell,\iota)}\}, \mathbf{A}' \leftarrow \mathbf{A}^{(\ell,\iota)}$
- 13: **return** $\{\mathbf{P}^{(\ell)}\}, \mathbf{A}'$

4.3. Nearest vector method to find $\mathbf{A}^\star$

Upon obtaining $\mathbf{A}' = [\mathbf{a}'_1, \dots, \mathbf{a}'_{M_r}]$, which is the optimal solution for $\mathbf{A} \in \mathbb{A}_c$, the next step in the BCD-AP algorithm is to apply AP for projecting $\mathbf{a}'_{m_r}$ onto $\mathbb{A}_r$. The nearest element of $\mathbf{a}'_{m_r}$ in $\mathbb{A}_r$ for all m_r in the following problem

$$\underset{\mathbf{a}_{m_r}, \forall m_r}{\text{minimize}} \|\mathbf{a}_{m_r} - \mathbf{a}'_{m_r}\|_2^2 \text{ subject to (33f) and (33g),} \quad (76)$$

yields $\mathbf{a}_{m_r}^\star$, the m_r -th column of $\mathbf{A}^\star$. This is effectively a matrix nearness problem with specified column norms and PARs. It arises in structured tight frame design problems and is solved via AP [37, 36]. Using "nearest vector with low PAR" algorithm [37], we find $\mathbf{a}_{m_r}^{(\ell)}$ for all m_r recursively at the ℓ -th iteration of the BCD-AP MRMC algorithm (see Algorithm 3).

Algorithm 3 Nearest vector method to find $\mathbf{A}^{(\ell)}$

Input: $\mathbf{A}' = [\mathbf{a}'_1, \dots, \mathbf{a}'_{M_r}]$, P_{r,m_r} , γ_{m_r} , $\forall m_r$
Output: $\mathbf{A}^{(\ell)} = [\mathbf{a}^{(\ell)}_1, \dots, \mathbf{a}^{(\ell)}_{M_r}]$

- 1: **for** $m_r = 1, \dots, M_r$ **do**
- 2: Normalize $\mathbf{a}'_{m_r}$ to unit norm; define $\sigma_{m_r} = \sqrt{P_{r,m_r}\gamma_{m_r}/K}$
- 3: $P \leftarrow$ number of elements in $\mathbf{a}'_{m_r}$ with the least magnitude
- 4: $\varpi \leftarrow$ indices of the elements in $\mathbf{a}'_{m_r}$ with the least magnitude
- 5: **if** $\min(|\mathbf{a}'_{m_r}[k]|) = 0, \forall k \in \varpi$ **then**
- 6:
$$\mathbf{a}^{(\ell)}_{m_r}[k] = \begin{cases} \sqrt{\frac{P_{r,m_r}(K-P)\sigma_{m_r}^2}{P}} & \text{if } k \in \varpi, \\ \sigma_{m_r} e^{j\angle \mathbf{a}'_{m_r}[k]} & \text{if } k \notin \varpi \end{cases}$$
- 7: **else** $\rho = \sqrt{\frac{P_{r,m_r}(K-P)\sigma_{m_r}^2}{\sum_{k \in \varpi} |\mathbf{a}'_{m_r}[k]|^2}}$ and
- 8:
$$\mathbf{a}^{(\ell)}_{m_r}[k] = \begin{cases} \rho \mathbf{a}'_{m_r}[k] & \text{if } k \in \varpi, \\ \sigma_{m_r} e^{j\angle \mathbf{a}'_{m_r}[k]} & \text{if } k \notin \varpi \end{cases}$$
- 9: **return** $\mathbf{A}^{(\ell)} = [\mathbf{a}^{(\ell)}_1, \dots, \mathbf{a}^{(\ell)}_{M_r}]$

Once $\{\mathbf{P}^{(\ell)}\}$ and $\mathbf{A}^{(\ell)}$ are known, we update $\{\mathbf{U}^{(\ell)}\}$ with WMMSE solutions from Section 4.1. Algorithm 4 summarizes the BCD-AP MRMC with $\ell_{\max}$ the maximum number of iterations.

Algorithm 4 BCD-AP MRMC algorithm

Input: $\mathbf{A}$, $\ell_{\max}$, $l_{\max}$, $t_{u,\max}$, $t_{d,\max}$
Output: Optimal UL/DL precoders $\{\mathbf{P}^*\}$, MIMO radar code matrix $\mathbf{A}^*$, and LRFs $\{\mathbf{U}^*\}$

- 1: Initialize $\{\mathbf{P}^{(0)}\} \triangleq \{\mathbf{P}^{(0)}_{u,i}[k], \mathbf{P}^{(0)}_{d,j}[k], \forall \{i, j, k\}\}$ and $\mathbf{A}^{(0)} = [\mathbf{a}^{(0)}[0]; \dots; \mathbf{a}^{(0)}[K]]$
- 2: $\{\mathbf{U}^{(0)}\} \xleftarrow{(43),(44),(45)} \{\mathbf{P}^{(0)}\}$ and $\mathbf{A}^{(0)}$
- 3: Set the alternating projection iteration index $\ell = 0$
- 4: **repeat**
- 5: $\{\mathbf{P}^{(\ell+1)}\}, \mathbf{A}' \xleftarrow[l_{\max}, t_{u,\max}, t_{d,\max}]{\text{Algorithm 2}} \{\mathbf{P}^{(\ell)}\}, \mathbf{A}^{(\ell)}, \{\mathbf{U}^{(\ell)}\}$
- 6: $\mathbf{A}^{(\ell+1)} \xleftarrow{\text{Algorithm 3}} \mathbf{A}'$
- 7: $\{\mathbf{U}^{(\ell+1)}\} \xleftarrow{(43),(44),(45)} \{\mathbf{P}^{(\ell+1)}\}$ and $\mathbf{A}^{(\ell+1)}$
- 8: $\ell \leftarrow \ell + 1$
- 9: **until** $\ell > \ell_{\max}$
- 10: $\{\mathbf{P}^*\} \leftarrow \{\mathbf{P}^{(\ell)}\}$, $\mathbf{A}^* \leftarrow \mathbf{A}^{(\ell)}$, and $\{\mathbf{U}^*\} \leftarrow \{\mathbf{U}^{(\ell)}\}$
- 11: **return** $\{\mathbf{P}^*\}, \mathbf{A}^*, \{\mathbf{U}^*\}$

4.4. Complexity and Convergence

The subgradient algorithms are guaranteed to converge to λ^* and μ^* as long as $\beta_{u,i}^{(t)}[k]$, $\beta_d^{(t)}[k]$, $\epsilon_{u,i}^{(t)}[k]$, and $\epsilon_{d,j}^{(t)}[k]$ are sufficiently small; their computational complexities are $\mathcal{O}(I)$ and $\mathcal{O}(J)$, respectively [30]. The WMMSE-MRMC algorithm converges locally because its alternating procedure produces a monotonically non-increasing sequence of

iterates, $\{\Xi_{\text{wmmse}}^{(\ell,i)}\}$; see Appendix C of [38] for proof. However, Ξ_{wmmse} is not jointly convex on $\{\mathbf{P}^{(\ell)}\}$ and $\mathbf{A}^{(\ell)}$. Hence, the global convergence is not guaranteed [38, 28]. The computational complexities of WMMSE-MRMC to update $\mathbf{P}_{u,i}[k]$, $\mathbf{P}_{d,j}[k]$, and $\mathbf{a}[k]$ at each iteration are, respectively, $\mathcal{O}\left((N_i^u D_i^u)^3\right)$, $\mathcal{O}\left((M_c D_j^d)^3\right)$, and $\mathcal{O}(M_r^3)$, primarily because of complexity in solving Sylvester equations. The total per-frame complexity of each iteration for Algorithm 2 to solve (58) is $\mathcal{O}\left(I^2 (N_i^u D_i^u)^3\right) + \mathcal{O}\left(J^2 (M_c D_j^d)^3\right) + \mathcal{O}(N_r^2 M_r^3)$. GB search on the other hand, is more computationally complex with $\mathcal{O}\left(I^3 (N_i^u D_i^u)^3\right)$ per iteration multiplications for DL [27]. The objective function in (76) satisfies the Kurdyka-Łojasiewicz property. Therefore, the sequence $\{\mathbf{a}_{m_r}^{(\ell)}\}$ generated by Algorithm 3 at the ℓ -th step is convergent for all m_r [36]. The BCD-AP MRMC algorithm also converges to the local optimum with a convergence rate of $\mathcal{O}(1/\ell_{\max})$ [31]. In general, initialization methods affect the convergence rate and local optimal values of BCD-AP MRMC. As suggested in [28, 38], to reasonably *approach* the global optimum, one can perform random precoder initializations and average over a large number of channel realizations while keeping track of the best result.

5. Numerical Experiments

We validated our spectral co-design approach through extensive numerical experiments. Throughout this section, we assume the noise variances $\sigma_r^2 = \sigma_B^2 = \sigma_d^2 = 0.001$. We assume unit small scale fading channel gains, namely, the elements of $\mathbf{H}_{B,j}$, $\mathbf{H}_{i,B}$, $\mathbf{H}_{i,j}$, α_{Bm,n_r} , and α_{i,n_r} are drawn from $\mathcal{CN}(0, 1)$. We model the self-interfering channel $\mathbf{H}_{BB}$ as $\mathcal{CN}\left(\sqrt{\frac{\sigma_{\text{SI}}^2 K_B}{1+K_B}} \hat{\mathbf{H}}_{BB}, \frac{\sigma_{\text{SI}}^2}{1+K_B} \mathbf{I}_{N_c} \otimes \mathbf{I}_{M_c}\right)$, where σ_{SI} is the SI attenuation coefficient that characterizes the effectiveness of SI cancellation [27], the Rician factor $K_B = 1$, and $\hat{\mathbf{H}}_{BB} \in \mathbb{C}^{N_c \times M_c}$ is an all-one matrix [28]. Define the signal-to-noise ratios (SNRs) associated with the MIMO radar, DL, and UL as $\text{SNR}_r = P_{r,m_r}/\sigma_r^2$, $\text{SNR}_{\text{DL}} = P_B/\sigma_d^2$, and $\text{SNR}_{\text{UL}} = P_{u,i}/\sigma_0^2$ [38]. The clutter power $\sigma_c^2 = \sigma_{m_r c n_r}^2$ for all m_r and n_r and clutter-to-noise ratio (CNR) is $\text{CNR} = \sigma_c^2/\sigma_0^2$. Then, together with the direct path components, they are received at the IBFD communications Rxs. We model $\mathbf{h}_{m_r,B}$ and $\mathbf{h}_{m_r,j}$ as $\mathcal{CN}\left(\sqrt{\frac{1}{\kappa+1}} \boldsymbol{\mu}_{m_r,B}, \frac{\eta_{m_r,B}^2}{\kappa+1} \mathbf{I}_{N_c}\right)$, and $\mathcal{CN}\left(\sqrt{\frac{1}{\kappa+1}} \boldsymbol{\mu}_{m_r,j}, \frac{\eta_{m_r,j}^2}{\kappa+1} \mathbf{I}_{N_j^d}\right)$, where $\kappa = 1$, $\boldsymbol{\mu}_{m_r,B} = 0.11 \mathbf{1}_{N_c}$, $\boldsymbol{\mu}_{m_r,j} = 0.051 \mathbf{1}_{N_j^d}$, $\eta_{m_r,B}^2 = 0.3$, $\eta_{m_r,j}^2 = 0.5$.

Unless otherwise stated, we use the following parameter values: number of radar TxS and RxS: $M_r = N_r = 4$; number of communications Tx and Rx antennas: $M_c = N_c = 4$; $I = J = 2$; $N_i^u = d_i^u = N_j^d = d_j^d = 2$, for all $\{i, j\}$; $\text{SNR}_{\text{DL}} = \text{SNR}_{\text{UL}} = 10$ dB; $\sigma_{\text{SI}}^2 = 0$ dB; $\text{CNR} = 20$ dB; radar PAR $\gamma_{m_r} = 3$ dB; number of communications frames or radar PRIs $K = 8$; number of symbols in each frame or range cells in each radar PRI $N = 32$; radar CUT index $n_t = 4$; UL (DL) indices of interest $n_{\text{TB}} = 2$ ($n_{\text{r,d}} = 3$); QoS of UL (DL): $R_u = \log_2(1 + \frac{\text{SNR}_{\text{UL}}}{M_r * \text{SNR}_r + \text{SNR}_{\text{DL}} + (I-1) * \text{SNR}_{\text{UL}}})$ bits/s/Hz ($R_d = \log_2(1 + \frac{\text{SNR}_{\text{DL}}/J}{M_r * \text{SNR}_r + \text{SNR}_{\text{DL}} * (J-1)/J + I * \text{SNR}_{\text{UL}}})$ bits/s/Hz); The normalized Doppler shifts $f_{m_r t_r} T_r$ and $f_{B t_r} T_r$ are uniformly distributed in $[0.05, 0.325]$ for each channel realization [23]. The numbers of iterations for the subgradient, weighted minimum mean-squared-error (WMMSE)-MRMC, and BCD-AP MRMC algorithms are $t_{u,\max} = t_{d,\max} = 200$, $t_{\max} = 1$, $\ell_{\max} = 2000$. We use uniform weights $\alpha_i^u = \alpha_j^d = \alpha_{n_r}^r = \frac{1}{(I+J+N_r)}$ for all $\{n_r, i, j\}$.

5.1. Convergence Analysis

We demonstrate the convergence of the BCD-AP MRMC algorithm with different initialization settings for the FD communications precoders $\{\mathbf{P}^{(0)}\}$ with SNR_r equal to -5 dB, 0 dB, 5 dB, and 10 dB. The first setting, dubbed as the deterministic initialization, initializes $\mathbf{P}_{d,j}^{(0)}[k]$ as the first D_j^d columns of the right singular matrix of $\mathbf{H}_{B,j}$ and $\mathbf{P}_{u,i}^{(0)}[k]$ as the first D_j^d columns of the right singular matrix of $\mathbf{H}_{i,B}$. Then, scale the non-zero singular values of $\mathbf{P}_{d,j}^{(0)}[k]$ and $\mathbf{P}_{u,i}^{(0)}[k]$ to be $P_B/J D_j^d$ and P_U/D_i^u , respectively, for all $\{i, j, k\}$ [38]. The second method, or the random initialization, generates the singular vectors of $\mathbf{P}_{u,i}^{(0)}[k]$ and $\mathbf{P}_{d,j}^{(0)}[k]$ as two random matrices drawn from $\mathcal{CN}(0, 1)$ and normalizes singular values in the same way as the deterministic initialization. The first initialization method offers lower computational complexity

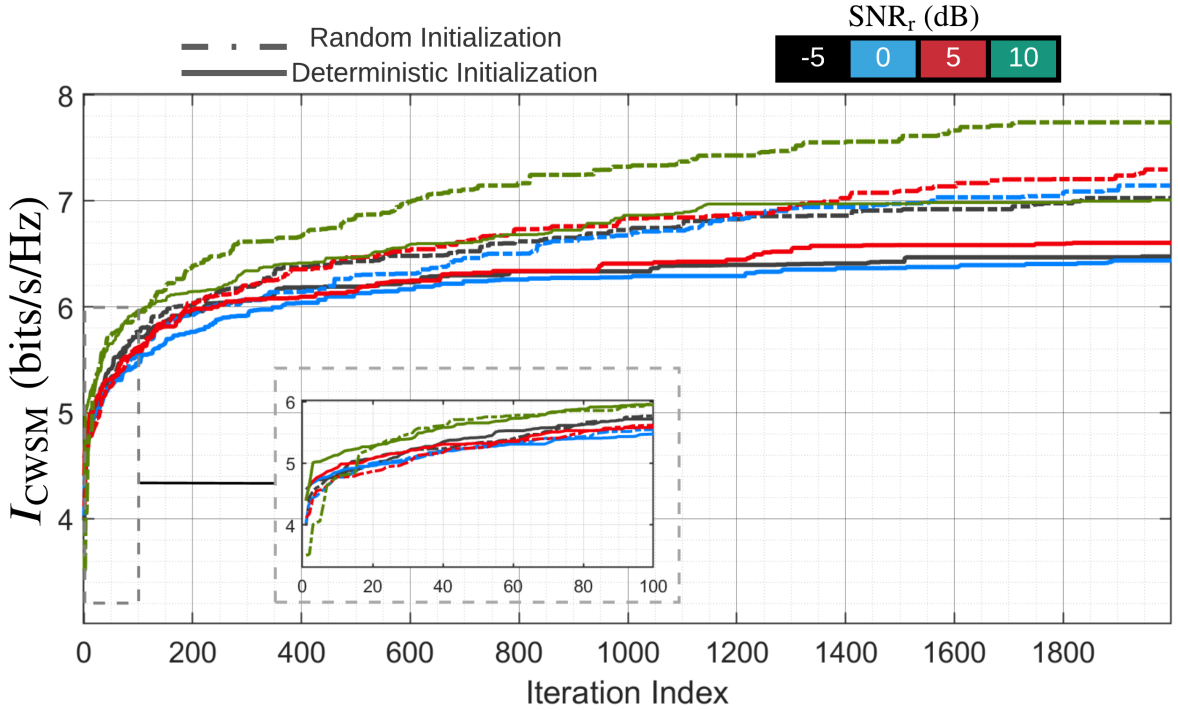


Figure 2: Convergence behaviors of the BCD-AP MRMC algorithm with two initialization methods and multiple SNR_r s.

than the latter. Figure 2 shows both initialization approaches achieve convergence as the number of iterations increases for all the simulated radar SNR values. However, the proposed algorithm with random initialization consistently outperforms its counterpart with the deterministic initialization at the cost of computational complexity. The inset plot in Figure 2 numerically demonstrates the efficiency of the proposed WMMSE-based algorithm; it converges within 80 iterations for both deterministic and random initializations.

5.2. Joint Radar-Communications Performance

We evaluated the co-design performance by observing the mutual impact of the statistical MIMO radar and IBFD MU-MIMO communications on each other. Figure 3 compares the comprehensive performance of the proposed BCD-AP MRMC algorithm with the existing communications precoding schemes and radar codes when the radar SNR changes³ with perfect and imperfect CSI. We set η_{CSI}^2 to 0.1. The proposed co-design algorithm demonstrates a significant advantage over other precoding and radar code matrices for radar SNRs between -20 dB and 20 dB, with and without CIR errors. It is noteworthy that our approach with imperfect CSI yields a higher I_{CWSM} than even the perfect CSI scenarios of conventional techniques. This is explained by the fact that local CSI for each Tx, i.e., the channel coefficients that are directly connected to this Tx, is needed at each iteration of Algorithm 2 [27].

We further investigated the impacts of the presence of clutter on the IBFD MU-MIMO communications in Figure 4a and different numbers of communications UEs on the MIMO radar target detection in Figure 4b. We plot the total MI of the IBFD MU-MIMO system $I_{\text{FD}} = \sum_{k=0}^{K-1} \left[\sum_{i=1}^I \alpha_i^u I_i^u[k] + \sum_{j=1}^J \alpha_j^d I_j^d[k] \right]$ versus CNRs in Figure 4a, where I_{FD} remains relatively invariant when CNR increases from 10 to 40 dB and the proposed algorithm outperforms other precoding strategies for simulated CNRs. We performed 10^4 Monte Carlo simulations with $D_i^u = 2$ (1), $D_j^d = 1$ (2) for UL (DL) ROC curves in Fig. 4b. The number of DL UEs is 2 for the dashed curves, and the number of UL UEs is 2

³Note that the proposed UL precoding algorithm is applied to BD and NSP; the proposed DL precoding algorithm is applied to the uniform UL precoding case, and the proposed UL and DL precoding algorithms are applied to the cases of random and uncoded radar codes.

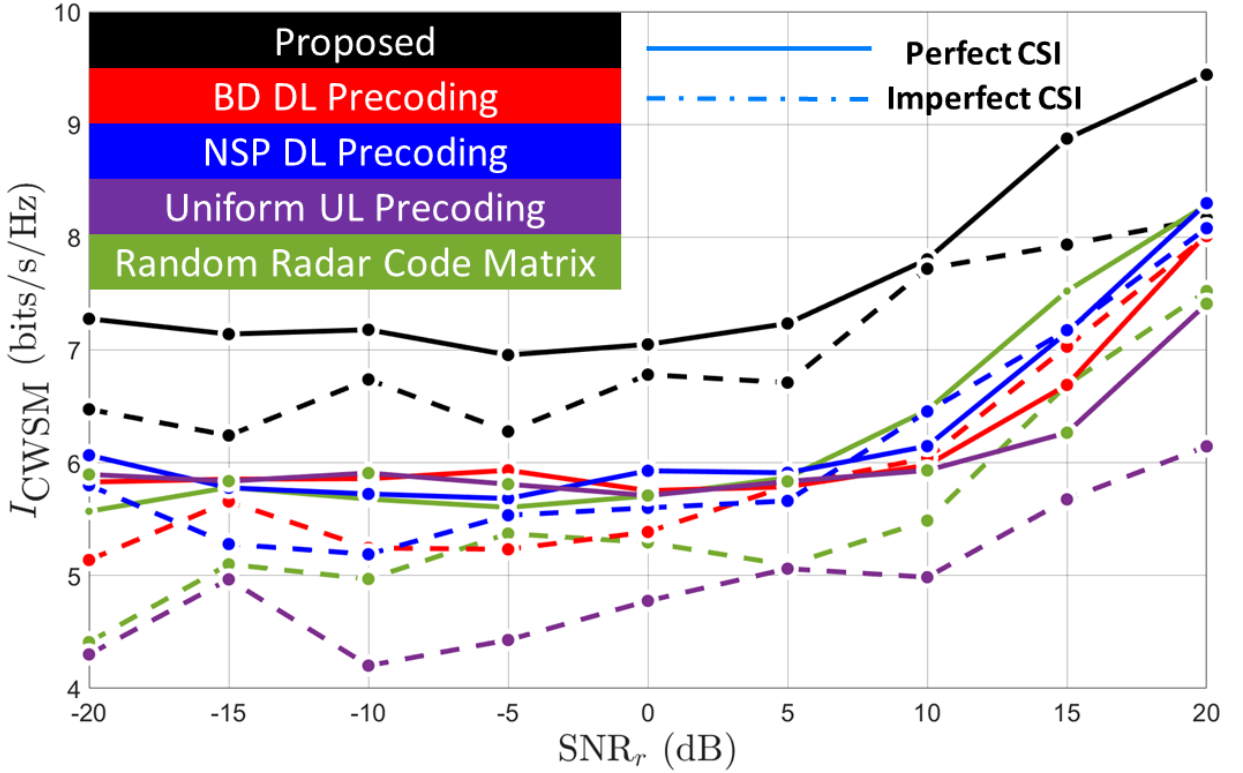


Figure 3: Proposed co-design compared with the conventional communications precoding and radar coding techniques under varying radar SNRs with and without CSI errors.

for the solid curves, which shows that the number of UL UEs is negatively correlated with the MIMO radar detection performance while the cooperation between the DL and the MIMO radar sustains the radar detection performance. Figure 5 depicts the impact of FD SI attenuation level ranging from -30 to 0 dB⁴ on the joint radar-communications co-design measured via I_{CWSM} . As expected, the stronger the SI cancellation, the better the MRMC system outputs. Our proposed algorithm outperforms other precoders even at very low SI cancellation or high values of σ_{SI}^2 .

6. Summary

We jointly designed UL/DL precoders, MIMO radar code matrix, and LRFs for both statistical MIMO radar and IBFD MU-MIMO communications operating in the same frequency bands. We proposed BCD-AP MRMC algorithm that ensures convergence and, as shown in the previous companion paper (Part I) [19], delivers performance benefits for both the radar and communications systems. This paper, through extensive experimentation, demonstrated the rapid convergence of the BCD-AP MRMC algorithm using two different initialization schemes. Our co-designed DL and radar systems exhibited robustness against substantial UL interference. Moreover, our optimized precoders and radar codes maintained stable DL and UL data rates as the CNR increased. Furthermore, we validated the robustness of our methods against imperfect CIR estimates and a wide range of SI attenuation levels. In summary, our proposed spectral co-design framework and the BCD-AP MRMC algorithm offer significant performance improvements for both the MIMO radar and IBFD MU-MIMO communications system. The following companion paper (Part III) [40] considers distributed beamforming and tracking aspects of the same system.

⁴Note that the NSP performance at $\sigma_{SI}^2 = 0$ dB is different than Figure 6 in the companion paper Part I, which employs random initialization. The experiment in this paper used the deterministic initialization.

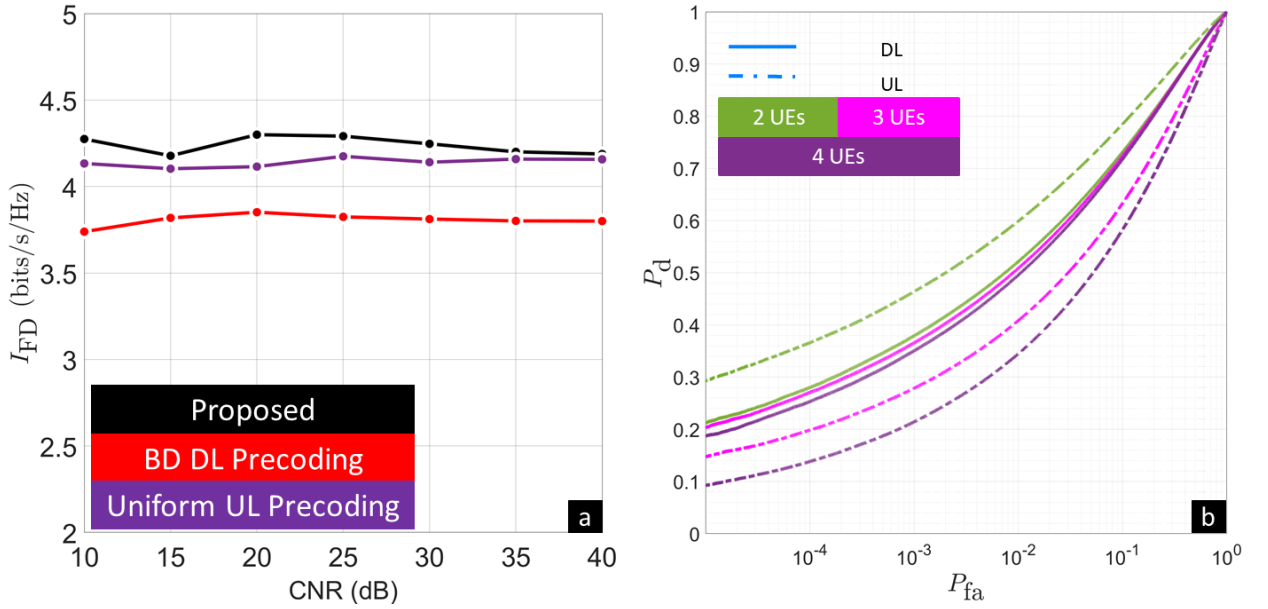


Figure 4: Joint radar and communications analyses: (a) IBFD MU-MIMO performance versus CNRs. (b) ROC curves with varying numbers of UL/DL UEs.

A. Derivation of Gradients

Denote the complex gradient operator for a scalar real-valued function with a complex-valued matrix argument $f(\mathbf{Z}, \mathbf{Z}^*)$ as $\nabla_{\mathbf{Z}} f = \frac{\partial f}{\partial \mathbf{Z}^*}$. From the derivative formula $\frac{\partial}{\partial \mathbf{X}^*} \text{Tr}(\mathbf{B}^T \mathbf{X}^\dagger \mathbf{C} \mathbf{X} \mathbf{B}) = \mathbf{C}^T \mathbf{X} \mathbf{B} \mathbf{B}^T + \mathbf{C} \mathbf{X} \mathbf{B} \mathbf{B}^T$ [27], the gradients of Ξ_{UL} and Ξ_{DL} w.r.t. $\mathbf{P}_{u,i}[k]$, $\mathbf{P}_{d,j}[k]$ and $\mathbf{a}[k]$ are

$$\begin{aligned} \nabla_{\mathbf{P}_{u,i}[k]} \Xi_{UL} &= 2\mathbf{H}_{i,B}^\dagger \xi_{UL} \mathbf{H}_{i,B} \mathbf{P}_{u,i}[k] - 2\alpha_i^u \mathbf{H}_{i,B}^\dagger \mathbf{U}_{u,i}^\dagger[k] \mathbf{W}_{u,i}[k], \\ \nabla_{\mathbf{P}_{d,j}[k]} \Xi_{DL} &= 2 \sum_{g=1}^J \mathbf{H}_{B,g}^\dagger \xi_{d,g} \mathbf{H}_{B,g} \mathbf{P}_{d,j}[k] - 2\alpha_j^d \mathbf{H}_{B,j}^\dagger \mathbf{U}_{d,j}^\dagger[k] \mathbf{W}_{d,j}[k], \\ \nabla_{\mathbf{P}_{d,j}[k]} \Xi_{UL} &= 2\mathbf{H}_{BB}^\dagger \xi_{UL}[k] \mathbf{H}_{BB} \mathbf{P}_{d,j}[k], \\ \nabla_{\mathbf{P}_{u,i}[k]} \Xi_{DL} &= 2 \sum_{g=1}^J \mathbf{H}_{i,g}^\dagger \xi_{d,g}[k] \mathbf{H}_{i,g} \mathbf{P}_{u,i}[k], \\ \nabla_{\mathbf{a}[k]} \Xi_{UL} &= 2\mathbf{H}_{rB}^\dagger \xi_{UL}[k] \mathbf{H}_{rB} \mathbf{a}[k], \end{aligned} \quad (77)$$

and

$$\nabla_{\mathbf{a}[k]} \Xi_{DL} = 2 \sum_{g=1}^J \mathbf{H}_{r,g}^\dagger \xi_{d,g}[k] \mathbf{H}_{r,g} \mathbf{a}[k], \quad (78)$$

respectively, where $\xi_{UL}[k] = \sum_{q=1}^I \alpha_q^u \mathbf{U}_{u,q}^\dagger[k] \mathbf{W}_{u,q}[k] \mathbf{U}_{u,q}[k]$ and $\xi_{d,g}[k] = \alpha_g^d \mathbf{U}_{d,g}^\dagger[k] \mathbf{W}_{d,g}[k] \mathbf{U}_{d,g}[k]$. The gradients of Ξ_r w.r.t. $\mathbf{P}_{u,i}[k]$, $\mathbf{P}_{d,j}[k]$, and $\mathbf{a}[k]$ are respectively shown as

$$\nabla_{\mathbf{P}_{u,i}[k]} \Xi_r = 2 \sum_{n_r=1}^{N_r} \sum_{m=1}^K \text{Re} \left(\xi_{r,n_r}(k, m) \Sigma_{i,n_r}^{(m,k)} \right) \mathbf{P}_{u,i}[m] \mathbf{d}_{u,i}[m, n_t - n_u] \mathbf{d}_{u,i}^\dagger[k, n_t - n_u]$$

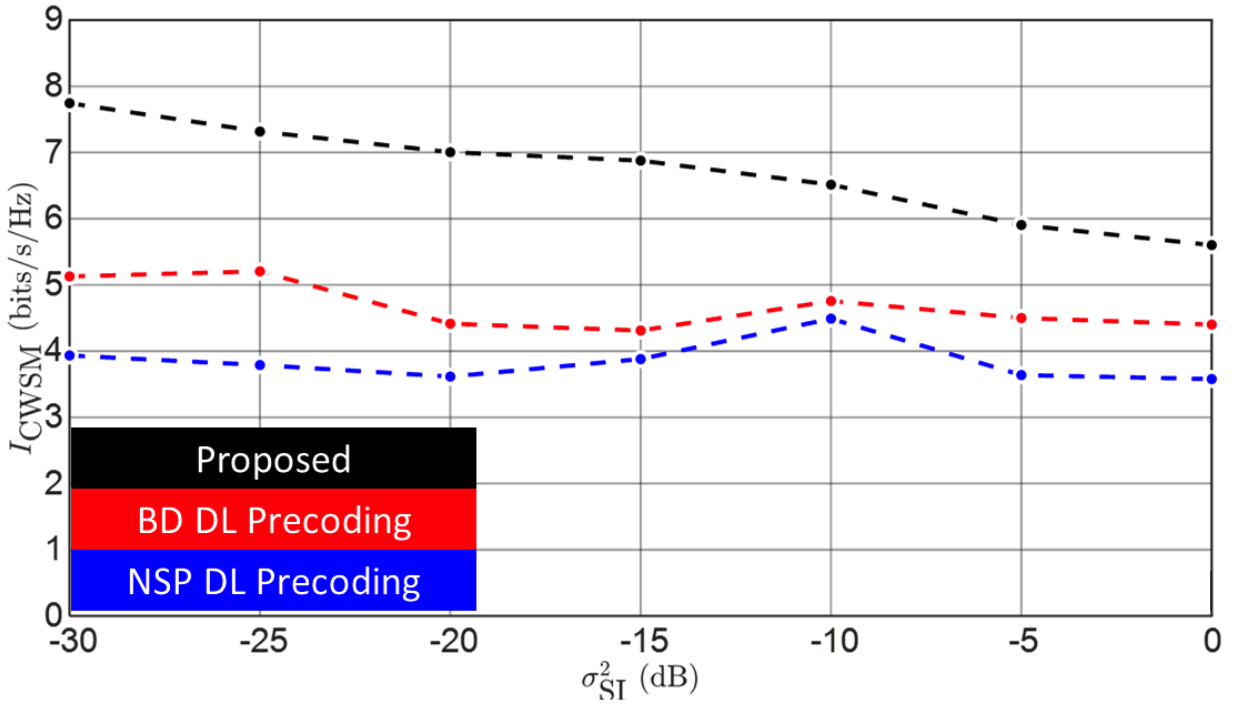


Figure 5: Impact of the FD SI on the proposed radar-communications co-design.

$$\begin{aligned}
\nabla_{\mathbf{p}_{d,j}[k]} \Xi_r &= 2 \sum_{n_r=1}^{N_r} \sum_{m=1}^K \text{Re} \left(\xi_{r,n_r}(k, m) \Sigma_{\text{Bt},n_r}^{(m,k)} \right) \sum_{g=1}^J \left\{ \mathbf{p}_{d,g}[m] \mathbf{d}_{d,g}[m, 0] \mathbf{d}_{d,j}^\dagger[k, 0] \right\} \\
&+ 2 \sum_{n_r=1}^{N_r} \sum_{m=1}^K \text{Re} \left(\xi_{r,n_r}(k, m) \Sigma_{\text{Bm},n_r}^{(m,k)} \right) \sum_{g=1}^J \mathbf{p}_{d,g}[m] \mathbf{d}_{d,g}[m, n_t - n_{\text{Bm}}] \mathbf{d}_{d,j}^\dagger[k, n_t - n_{\text{Bm}}] \\
&- 2 \sum_{n_r=1}^{N_r} \sum_{m=1}^K \text{Re} \left(\Sigma_{\text{Bt},n_r}^{(m,k)} \right) \mathbf{J}_B^\top \mathbf{J}_h^\top[m] \mathbf{W}_{r,n_r} \mathbf{u}_{r,n_r}[k] \mathbf{d}_{d,j}^\dagger[k, 0], \\
\nabla_{\mathbf{a}[k]} \Xi_r &= -2 \sum_{n_r=1}^{N_r} \sum_{m=1}^K \text{Re} \left(\Sigma_{\text{rt},n_r}^{(m,k)} \right) \mathbf{J}_r^\top \mathbf{J}_h^\top[m] \mathbf{W}_{r,n_r} \mathbf{u}_{r,n_r}[k] \\
&+ 2 \sum_{n_r=1}^{N_r} \sum_{m=1}^K \left[\text{Re} \left(\xi_{r,n_r}(k, m) \Sigma_{\text{rt},n_r}^{(m,k)} \right) + \text{Re} \left(\xi_{r,n_r}(k, m) \Sigma_{\text{c},n_r} \right) \right] \mathbf{a}[m],
\end{aligned}$$

where $\xi_{r,n_r}(k, m) = \alpha_{n_r}^\top \text{Tr} \left\{ \mathbf{u}_{r,n_r}^\dagger[k] \mathbf{W}_{r,n_r} \mathbf{u}_{r,n_r}[m] \right\}$. The derivatives of $f(\mathbf{X}, \mathbf{X}^*)$ in (69) w.r.t. $\mathbf{X}^*$ are thus approximated as $\frac{\partial f}{\partial \mathbf{X}^*} = \frac{\partial}{\partial \mathbf{X}_0^*} f(\mathbf{X}_0, \mathbf{X}_0^*)$. The chain rule for a scalar function $g(\mathbf{U}(\mathbf{X}, \mathbf{X}^*), \mathbf{U}^*(\mathbf{X}, \mathbf{X}^*))$ where g is dependent on $\mathbf{X}^*$ through the matrix $\mathbf{U}$ is [27]

$$\frac{\partial g}{\partial \mathbf{X}^*} = \frac{\text{Tr} \left\{ \left(\frac{\partial g}{\partial \mathbf{U}} \right)^\top \partial \mathbf{U} \right\}}{\partial \mathbf{X}^*} + \frac{\text{Tr} \left\{ \left(\frac{\partial g}{\partial \mathbf{U}^*} \right)^\top \partial \mathbf{U}^* \right\}}{\partial \mathbf{X}^*}. \quad (79)$$

With (79) and $\partial \log|\mathbf{X}| = \text{Tr} \{ \mathbf{X}^{-1} \partial \mathbf{X} \}$ [27], the derivatives of $R_{u,i}[k]$ and $R_{d,j}[k]$ w.r.t. $\mathbf{P}_{u,i}[k]$, $\mathbf{P}_{d,j}[k]$ based on their associated first order Taylor series expansions are

$$\begin{aligned}\nabla_{\mathbf{P}_{u,i}[k]} R_{u,i}[k] &= \mathbf{H}_{i,B}^\dagger \left(\mathbf{R}_{u,i}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{i,B} \tilde{\mathbf{P}}_{u,i}[k] \tilde{\mathbf{E}}_{u,i}^*[k], \\ \nabla_{\mathbf{P}_{d,j}[k]} R_{d,j}[k] &= \mathbf{H}_{B,j}^\dagger \left(\mathbf{R}_{d,j}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{B,j} \tilde{\mathbf{P}}_{d,j}[k] \tilde{\mathbf{E}}_{d,j}^*[k], \\ \nabla_{\mathbf{P}_{u,i}[k]} R_{u,q}[k] &= -\mathbf{H}_{i,B}^\dagger \left(\mathbf{R}_{u,q}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{q,B} \mathbf{P}_{u,q}[k] \mathbf{E}_{u,q}^*[k] \times \mathbf{P}_{u,q}^\dagger[k] \mathbf{H}_{q,B}^\dagger \left(\mathbf{R}_{u,q}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{i,B} \tilde{\mathbf{P}}_{u,i}[k], \quad q \neq i, \\ \nabla_{\mathbf{P}_{d,j}[k]} R_{u,i}[k] &= -\mathbf{H}_{B,B} \left(\mathbf{R}_{u,i}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{i,B}^\dagger \mathbf{P}_{u,i}[k] \mathbf{E}_{u,i}^*[k] \mathbf{H}_{i,B}^\dagger \mathbf{P}_{u,i}^\dagger[k] \left(\mathbf{R}_{u,i}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{B,B} \tilde{\mathbf{P}}_{d,j}, \\ \nabla_{\mathbf{P}_{u,i}[k]} R_{d,j}[k] &= -\mathbf{H}_{i,j}^\dagger \left(\mathbf{R}_{d,j}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{B,j} \mathbf{P}_{d,j}[k] \mathbf{E}_{d,j}^*[k] \mathbf{P}_{d,j}^\dagger[k] \mathbf{H}_{B,j}^\dagger \left(\mathbf{R}_{d,j}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{i,j} \tilde{\mathbf{P}}_{u,i}[k], \\ \nabla_{\mathbf{P}_{d,j}[k]} R_{d,g}[k] &= -\mathbf{H}_{B,g}^\dagger \left(\mathbf{R}_{d,g}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{B,g} \mathbf{P}_{d,g}[k] \mathbf{E}_{d,g}^*[k] \mathbf{P}_{d,g}^\dagger[k] \mathbf{H}_{B,g}^\dagger \left(\mathbf{R}_{d,g}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{B,g} \tilde{\mathbf{P}}_{d,j}[k], \quad g \neq j,\end{aligned}$$

where $\tilde{\mathbf{E}}_{u,i}^*[k] \triangleq \mathbf{E}_{u,i}^*[k] \left(\tilde{\mathbf{P}}_{u,i}[k] \right)$, $\tilde{\mathbf{E}}_{d,j}^*[k] = \mathbf{E}_{d,j}^*[k] \left(\tilde{\mathbf{P}}_{d,j}[k] \right)$. The derivatives of $R_{u,i}[k]$ and $R_{d,j}[k]$ w.r.t. $\mathbf{a}[k]$ are $\nabla_{\mathbf{a}[k]} R_{d,j}[k] = -\mathbf{H}_{r,j}^\dagger \left(\mathbf{R}_{d,j}^{\text{in}}[k] \right)^{-1} \mathbf{P}_{d,j}[k] \mathbf{H}_{B,j} \mathbf{E}_{d,j}^*[k] \mathbf{P}_{d,j}^\dagger[k] \mathbf{H}_{B,j}^\dagger \left(\mathbf{R}_{d,j}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{r,j} \mathbf{a}[k]$ and $\nabla_{\mathbf{a}[k]} R_{u,i}[k] = -\mathbf{H}_{rB}^\dagger \left(\mathbf{R}_{u,i}^{\text{in}}[k] \right)^{-1} \mathbf{P}_{u,i}[k] \mathbf{H}_{i,B} \mathbf{E}_{u,i}^*[k] \mathbf{P}_{u,i}^\dagger[k] \mathbf{H}_{i,B}^\dagger \left(\mathbf{R}_{u,i}^{\text{in}}[k] \right)^{-1} \mathbf{H}_{rB} \mathbf{a}[k]$.

References

- [1] K. V. Mishra, M. R. Bhavani Shankar, V. Koivunen, B. Ottersten, S. A. Vorobyov, Toward millimeter wave joint radar communications: A signal processing perspective, *IEEE Signal Processing Magazine* 36 (5) (2019) 100–114.
- [2] M. I. Skolnik, *Radar handbook*, 3rd Edition, McGraw-Hill, 2008.
- [3] B. Clerckx, C. Oestges, *MIMO wireless networks: Channels, techniques and standards for multi-antenna, multi-user and multi-cell systems*, Academic Press, 2013.
- [4] T. M. Cover, J. A. Thomas, *Elements of information theory*, 2nd Edition, John Wiley & Sons, 2006.
- [5] Y. Cui, V. Koivunen, X. Jing, Interference alignment based spectrum sharing for MIMO radar and communication systems, in: *IEEE International Workshop on Signal Processing Advances in Wireless Communications*, 2018, pp. 1–5.
- [6] A. Ayyar, K. V. Mishra, Robust communications-centric coexistence for turbo-coded OFDM with non-traditional radar interference models, in: *IEEE Radar Conference*, 2019, pp. 1–6.
- [7] S. H. Dokhanchi, B. S. Mysore, K. V. Mishra, B. Ottersten, A mmWave automotive joint radar-communications system, *IEEE Transactions on Aerospace and Electronic Systems* 55 (3) (2019) 1241–1260.
- [8] G. Duggal, S. Vishwakarma, K. V. Mishra, S. S. Ram, Doppler-resilient 802.11ad-based ultra-short range automotive joint radar-communications system, *IEEE Transactions on Aerospace and Electronic Systems* 56 (5) (2020) 4035–4048.
- [9] B. Li, A. Petropulu, Joint transmit designs for co-existence of MIMO wireless communications and sparse sensing radars in clutter, *IEEE Transactions on Aerospace and Electronic Systems* 53 (6) (2017) 2846–2864.
- [10] Q. He, Z. Wang, J. Hu, R. S. Blum, Performance gains from cooperative MIMO radar and MIMO communication systems, *IEEE Signal Processing Letters* 26 (1) (2019) 194–198.
- [11] A. M. Haimovich, R. S. Blum, L. J. Cimini, MIMO radar with widely separated antennas, *IEEE Signal Processing Magazine* 25 (1) (2008) 116–129.
- [12] E. Fishler, A. Haimovich, R. S. Blum, L. J. Cimini, D. Chizhik, R. A. Valenzuela, Spatial diversity in radars—models and detection performance, *IEEE Transactions on Signal Processing* 54 (3) (2006) 823–838.
- [13] J. Li, P. Stoica, MIMO radar with colocated antennas, *IEEE Signal Processing Magazine* 24 (5) (2007) 106–114.
- [14] S. Sun, Y. Hu, K. V. Mishra, A. P. Petropulu, Widely separated MIMO radar using matrix completion, *IEEE Transactions on Radar Systems* 2 (2024) 180–196.
- [15] M. Alaa-Kerahroodi, M. R. Bhavani Shankar, K. V. Mishra, B. Ottersten, Information theoretic approach for waveform design in coexisting MIMO radar and MIMO communications, in: *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2020, pp. 1–5.
- [16] S. H. Dokhanchi, M. R. Bhavani Shankar, K. V. Mishra, B. Ottersten, Multi-constraint spectral co-design for colocated MIMO radar and MIMO communications, in: *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2020, pp. 4567–4571.
- [17] D. Bao, G. Qin, J. Cai, G. Liu, A precoding OFDM MIMO radar coexisting with a communication system, *IEEE Transactions on Aerospace and Electronic Systems* 55 (4) (2019) 1864–1877.
- [18] A. Khawar, A. Abdelhadi, C. Clancy, Target detection performance of spectrum sharing MIMO radars, *IEEE Sensors Journal* 15 (9) (2015) 4928–4940.
- [19] J. Liu, K. V. Mishra, M. Saquib, Co-designing statistical MIMO radar and in-band full-duplex multi-user MIMO communications – Part I: Signal processing, *Signal Processing Under review* (2024).

- [20] C. B. Barneto, S. D. Liyanaarachchi, M. Heino, T. Riihonen, M. Valkama, Full duplex radio/radar technology: The enabler for advanced joint communication and sensing, *IEEE Wireless Communications* 28 (1) (2021) 82–88.
- [21] A. Deligiannis, A. Daniyan, S. Lambotharan, J. A. Chambers, Secrecy rate optimizations for MIMO communication radar, *IEEE Transactions on Aerospace and Electronic Systems* 54 (5) (2018) 2481–2492.
- [22] S. Biswas, K. Singh, O. Taghizadeh, T. Ratnarajah, Coexistence of MIMO radar and FD MIMO cellular systems with QoS considerations, *IEEE Transactions on Wireless Communications* 17 (11) (2018) 7281–7294.
- [23] M. M. Naghsh, M. Modarres-Hashemi, M. A. Kerafroudi, E. H. M. Alian, An information theoretic approach to robust constrained code design for MIMO radars, *IEEE Transactions on Signal Processing* 65 (14) (2017) 3647–3661.
- [24] J. Liu, K. V. Mishra, M. Saquib, Transceiver co-design for full-duplex integrated sensing and communications, in: *IEEE Global Communications Conference*, 2022, pp. 3821–3826.
- [25] B. Tang, J. Tang, Y. Peng, MIMO radar waveform design in colored noise based on information theory, *IEEE Transactions on Signal Processing* 58 (9) (2010) 4684–4697.
- [26] X. Song, P. Willett, S. Zhou, P. B. Luh, The MIMO radar and jammer games, *IEEE Transactions on Signal Processing* 60 (2) (2012) 687–699.
- [27] A. C. Cirik, R. Wang, Y. Hua, M. Latva-aho, Weighted sum-rate maximization for full-duplex MIMO interference channels, *IEEE Transactions on Communications* 63 (3) (2015) 801–815.
- [28] P. Aquilina, A. C. Cirik, T. Ratnarajah, Weighted sum rate maximization in full-duplex multi-user multi-cell MIMO networks, *IEEE Transactions on Communications* 65 (4) (2017) 1590–1608.
- [29] J. Liu, M. Saquib, Transmission design for a joint MIMO radar and MU-MIMO downlink communication system, in: *IEEE Global Conference on Signal and Information Processing*, 2018, pp. 196–200.
- [30] Wei Yu, R. Lui, Dual methods for nonconvex spectrum optimization of multicarrier systems, *IEEE Transactions on Communications* 54 (7) (2006) 1310–1322.
- [31] A. Beck, L. Tetruashvili, On the convergence of block coordinate descent type methods, *SIAM J. Optim.* 23 (2013) 2037–2060.
- [32] J. Liu, M. Saquib, Joint transmit-receive beamspace design for colocated MIMO radar in the presence of deliberate jammers, in: *Asilomar Conf. Signals Syst. Comput.*, 2017, pp. 1152–1156.
- [33] C. Chen, M. Li, X. Liu, Y. Ye, Extended ADMM and BCD for nonseparable convex minimization models with quadratic coupling terms: convergence analysis and insights, *Math. Program.* 173 (2019) 37–77.
- [34] Z. Zhang, M. Brand, Convergent block coordinate descent for training Tikhonov regularized deep neural networks, in: *Advances Neural Inf. Process. Syst.*, 2017, pp. 1721–1730.
- [35] E. Grossi, M. Lops, L. Venturino, Joint design of surveillance radar and MIMO communication in cluttered environments, *IEEE Transactions on Signal Processing* 68 (2020) 1544–1557.
- [36] Z. Zhu, X. Li, Convergence analysis of alternating projection method for nonconvex sets, *arXiv preprint arXiv:1802.03889v2* (2019).
- [37] J. A. Tropp, I. S. Dhillon, R. W. Heath, T. Strohmer, Designing structured tight frames via an alternating projection method, *IEEE Trans. Inf. Theory* 51 (1) (2005) 88–209.
- [38] Q. Shi, M. Razaviyayn, Z. Luo, C. He, An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel, *IEEE Transactions on Signal Processing* 59 (9) (2011) 4331–4340.
- [39] S. Boyd, J. Park, Subgradient methods (May 2014).
URL https://web.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf
- [40] S. Nayemuzzaman, J. Liu, K. V. Mishra, M. Saquib, Co-designing statistical MIMO radar and in-band full-duplex multi-user MIMO communications – Part III: Multi-target tracking, *Signal Processing* Under review (2024).