

Joint Transmit and Receive Beamforming Design in Full-Duplex Integrated Sensing and Communications

Ziang Liu, Sundar Aditya, *Member, IEEE*, Hongyu Li, *Student Member, IEEE*, and Bruno Clerckx, *Fellow, IEEE*

Abstract—Integrated sensing and communication (ISAC) has been envisioned as a solution to realize the sensing capability required for emerging applications in wireless networks. For a mono-static ISAC transceiver, as signal transmission durations are typically much longer than the radar echo round-trip times, the radar returns are drowned by the strong residual self interference (SI) from the transmitter, despite adopting sufficient SI cancellation techniques before digital domain - a phenomenon termed the *echo-miss* problem. A promising approach to tackle this problem involves the ISAC transceiver to be full-duplex (FD), and in this paper we jointly design the transmit and receive beamformers at the transceiver, transmit precoder at the uplink user, and receive combiner at the downlink user to simultaneously (a) maximize the uplink and downlink communication rate, (b) maximize the transmit and receive radar beampattern power at the target, and (c) suppress the residual SI. To solve this optimization problem, we proposed a penalty-based iterative algorithm. Numerical results illustrate that the proposed design can effectively achieve up to 60 dB digital-domain SI cancellation, a higher average sum-rate, and more accurate radar parameter estimation compared with previous ISAC FD studies.

Index Terms—Integrated sensing and communication, full-duplex, self-interference suppression, transmit/receive beamforming.

I. INTRODUCTION

Next generation wireless communication networks are expected to support high data transmission rates, high-quality wireless connectivity with massive devices, and highly accurate and robust sensing capability [1], [2]. To realize these requirements, integrated sensing and communication (ISAC), which unifies the signal processing procedures and hardware framework between sensing and communication systems, is widely viewed as a promising solution to efficiently utilize the available spectral, hardware and energy resources.

Challenge: *Echo miss*. In the existing ISAC literature, many studies assume that the sensing takes place in mono-static mode [3]–[7] due to its relative simplicity compared to other sensing configurations (e.g., multi-static). In a mono-static ISAC system, the transmit (TX)/receive (RX) antennas are co-located, resulting in the transmit (dual-function) waveform being known at the receiver. Hence, the receiver can use the transmit waveform as a reference waveform to extract

target information from the radar echo, thereby saving on the overhead associated with reference sharing. However, the transmit waveform is also expected to serve communications users in parallel, and typically communication frames are much longer than the radar echo round-trip times (RTTs). For example, in the 5G NR specifications [8], a standard radio frame has 10ms duration. For a target located at 100–1000m from the radar, its echo RTT is of the order of 1–10 μ s - orders of magnitude smaller than even the minimum unit of data scheduling (i.e., 1 slot = 0.5ms). Hence, the radar echo is drowned by the severe self interference (SI) from the transmitter, which causes receiver saturation, where the power of the received signal exceeds the analog-to-digital converter (ADC) dynamic range. Even if there is no ADC saturation due to the use of sufficient SI cancellation techniques before quantization, the radar echo may still be difficult to detect because it is masked by strong residual SI. We term this phenomenon the *echo-miss* problem. Thus, it is important to sufficiently suppress the residual SI to manageable levels.

To address the *echo-miss* problem and suppress the SI, one straightforward method is to physically separate the TX and RX antennas. The measurement-based study [9] shows that limited isolation capability can be achieved by a combination of directional isolation, absorber, and cross-polarization. Specifically, in the experiments, a 35cm separation between the TX and RX antennas, along with an absorber, was shown to realize 45dB passive suppression. However, the power level of the self interference (SI) can be large (i.e., up to 100dB larger than the receiver thermal noise floor [10], [11]). Thus, physical separation of TX and RX antennas may not entirely solve the *echo-miss* problem. Consequently, to integrate communication and sensing functions, the transceiver should work in the full-duplex (FD) mode to simultaneously transmit a dual-functional signal, receive the echo signal, and suppress the SI, caused by the leakage of the transmit signal to the receiver. Some attempts have been made for ISAC in the FD context, with [12] concentrating on waveform design by utilizing the waiting time of conventional pulsed radars to transmit communication signal in a single-antenna setup. For the multi-antenna case, [13] jointly optimizes relay beamformer, receive filter, and transmit power of the radar for a FD ISAC relay system, wherein the residual SI is assumed to be cancelled in advance by SI cancellation techniques. However, for the most part, the SI cancellation problem is underestimated by many ISAC studies [4], [5], [7], [14]–[16], which assume either ideal isolation between TX and RX or rely on the radar-function-

The authors are with the Communications & Signal Processing (CSP) Group at the Dept. of Electrical and Electronic Engg., Imperial College London, SW7 2AZ, UK. (e-mails: {ziang.liu20, s.aditya, c.li21, b.clerckx}@imperial.ac.uk).

focused SI cancellation method in [17].

Previous Approaches for SI Cancellation. The SI cancellation problem has been actively studied in FD wireless communication systems [18]–[20] (i.e., without the additional sensing functionality). In general, SI cancellation techniques can be adopted in the propagation [21], analog [22]–[24], and digital domains [25], [26]. As shown in Fig. 1, the propagation-domain cancellation aims to minimize the coupling between the transmit and receive direct paths. This kind of cancellation is achieved by techniques based on path loss, cross-polarization, and antenna directionality [9]. Beyond this, the analog-domain cancellation aims to suppress SI before the ADC, where a negative copy of the transmit waveform generated by the canceller circuit is subtracted from the received signal [20], [22], [23], [27]. Finally, as the last defense against SI, the digital-domain cancellation block utilizes either linear or non-linear adaptive filters to generate the negative of the residual SI [25], [26], and add it to the digital signal after the ADC.

The above SI cancellation techniques for communications rely on the uncorrelated nature between the SI (e.g., downlink data stream) and the signal of interest (SoI) (e.g., the uplink data stream); thus, the SI can be suppressed by adding its negative to the received signal without impairing the SoI. However, since the SoI in ISAC consists of uplink communications data and radar echoes that are correlated with the SI, it is challenging to effectively suppress the SI without distorting the radar echoes. To tackle this problem, utilizing the time-of-arrival (ToA) difference or the spatial angle-of-arrival (AoA) difference between the SI and echo are two promising approaches. With respect to the first approach, the early study [17] utilizes the temporal difference to generate a negative counterpart of the SI signal before the ADC (cf. Fig. 1 (b)), based on a gradient-learning method. Apart from adding this negative counterpart, an adaptive filter is also employed to generate a negative in digital domain to cancel the residual SI (cf. Fig. 1 (a)). In practice, many factors (e.g., RF taps, adaptive filter taps, and update algorithms) affect the accuracy of the generated negatives in both domains, which in turn, have a huge impact on the SI cancellation performance. If the SI signal is fast-changing, this approach may fail in tracking and can have high computational complexity.

In multi-antenna systems, an alternative way to suppress SI in ISAC is by employing the spatial AoA difference between the SI and SoI. In [28], the SI cancellation based on TX and RX beamforming design is adopted in the digital domain (cf. Fig. 1a). Specifically, the TX beamformer is the weighted sum of two separate beams probing at a downlink communication user and a radar target, whose power allocation is controlled by a parameter. In terms of the RX beamformer, the null-space projection (NSP) based on pseudoinverse operation is used to generate nulls in the angles of the downlink communication beam and SI. In [29] and [30], the NSP method is utilized to design hybrid TX and RX beamformer for sensing the target and communicating to a downlink user (cf. Figs. 1a and 1b). However, in these studies, the TX and RX beamformers are separately designed and only the RX beamformer is used for

TABLE I
NOVELTY COMPARISON WITH EXISTING FD ISAC LITERATURE

	Our work	[12]	[13]	[17]	[28]–[30]
Relay			✓		
Analog hardware design				✓	
Waveform design		✓			
Tx/Rx beamformers NSP design					✓
Tx/Rx beamformers joint design	✓		✓		
Uplink user	✓		✓		
Downlink user	✓	✓	✓		✓
Multi-antenna at users	✓				
Radar performance	✓	✓	✓	✓	✓
Communication performance	✓	✓			

SI cancellation. Recently, intelligent reflecting surfaces (IRS) have emerged as a means to boost the SoI, and thus reduce the effect of SI [31], [32], but this approach cannot actively cancel SI, and induces additional hardware and beamforming design complexity.

Thus, the potential of *joint ISAC TX-RX beamformer design* at transceiver to further suppress the SI has not been explored. In addition, the uplink communication performance in the FD ISAC system has not been analyzed. Given that the research on FD ISAC is still in its infancy (cf. Table I), we consider a mono-static FD ISAC multiple input multiple output (MIMO) system. In this system, (a) the TX-RX beamformers at the transceiver, (b) transmit precoder at the uplink user, and (c) receive combiner at the downlink user are jointly optimized. The aim of the optimization is to simultaneously (a) maximize the uplink and downlink rate, (b) maximize the transmit and receive radar beampattern power at the target, and (c) suppress the residual SI. Inspired by research adopting the penalty-dual decomposition (PDD) method (e.g., for FD mmWave hybrid beamforming [33], and IRS [34]), a penalty-based iterative algorithm is proposed to solve the optimization problem. Our proposed scheme can work when the received signal exceeds ADC dynamic range by adopting effective SI cancellation techniques before quantization.

Contributions and Overview of Results. In this paper, our contributions are summarized as follows:

- We first model a FD ISAC mono-static system to capture the *echo-miss* problem.
- To suppress the residual SI and preserve the two types of SoI (e.g., radar echo and uplink data), we formulate the joint TX-RX beamformer design problem for FD ISAC, where the objective function incorporates (a) uplink and downlink rates as the communications metric, (b) the transmit and receive radar beampattern power at a target as the sensing metric, and (c) the post-beamforming SI residual as a penalty term. Based on the equivalence between the rate maximization and the mean square error (MSE) minimization [35], [36], and inspired by the PDD method, an iterative algorithm with guaranteed convergence and acceptable complexity is developed using block coordination descent (BCD) methods. As seen in Table I, in contrast to the existing literature, our optimization framework concentrates on joint TX-RX beamformer design and directly cancel residual SI.

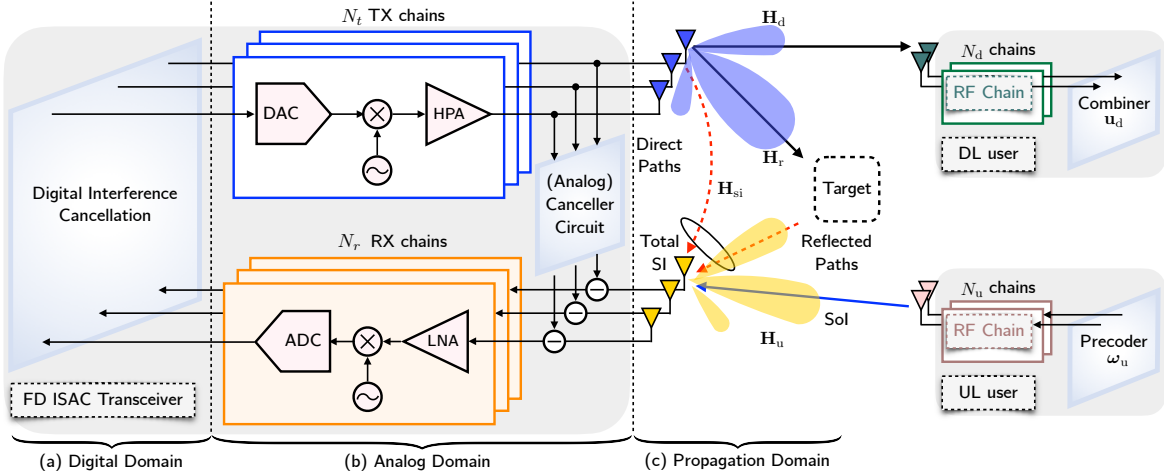


Fig. 1. Full-duplex integrated sensing and communication system, and illustration of the boundaries and contents of the SI suppression in the propagation, analog, and digital domains for multi-antenna configurations.

- The performance of the designed beamformers is validated via simulations, which show that up to 60dB residual SI can be effectively suppressed with better sum-rates for the uplink and downlink users and better radar parameter estimation performance (i.e., range-velocity and AoA detection) compared with the NSP method [28]–[30]. As shown in Table I, the performance of uplink communications, in particular, has not been thoroughly investigated in FD ISAC literature.

Organization of This Paper. The rest of the paper is organized as follows. The FD ISAC system model and the optimization framework for the joint ISAC TX-RX beamformers design is introduced in Section II. In Section III, the problem reformulation and the proposed joint ISAC TX-RX beamformers design algorithm are provided based on the BCD method. The convergence and complexity analysis of the proposed algorithm is given in Section IV, and numerical evaluations are presented in Section V. Finally, we conclude this work in Section VI.

Notation. The set of reals, integers, and complex numbers are denoted by \mathbb{R} , \mathbb{Z} , and \mathbb{C} , respectively. $\Re(x)$ and $\Im(x)$ denote the real and imaginary part of $x \in \mathbb{C}$, respectively. Continuous signals and discrete sequences are expressed by $x(t)$, $t \in \mathbb{R}$ and $x[k]$, $k \in \mathbb{Z}$, respectively. Matrices, vectors and scalars are written in capital boldface, small boldface and normal fonts, respectively. $[\mathbf{X}]_{i,:}$ and $[\mathbf{X}]_{:,j}$ denote the i -th row and j -th column of the matrix \mathbf{X} . $[\mathbf{X}]_{i,j}$ denotes the entry of the matrix \mathbf{X} at index (i, j) . Similarly, $[\mathbf{x}]_i$ for vector \mathbf{x} . \mathbf{X}^H , \mathbf{X}^T and \mathbf{X}^\dagger are used to denote conjugate-transpose, transpose and pseudo inverse of matrix \mathbf{X} , respectively. We use $\mathcal{E}(\cdot)$, $|\cdot|$, and $\|\cdot\|_2$ to denote statistical expectation, absolute value and Euclidean norm.

II. SYSTEM MODEL

A. Signal Model

As shown in Fig. 1, we consider a single-cell narrowband FD MIMO ISAC transceiver equipped with N_t transmit anten-

nas and N_r receive antennas. All antenna arrays are assumed to be uniform linear arrays (ULA) with half-wavelength spacing between adjacent antenna elements. The transceiver simultaneously serves one uplink user with N_u antennas, one downlink user with N_d antennas, and probes a target direction.

Let $s_d \in \mathbb{C}$ denote the ISAC downlink transmit symbol, and $s_u \in \mathbb{C}$ the uplink symbol. We assume both s_d and s_u have unit power. The received signal $\mathbf{y}_d \in \mathbb{C}^{N_d \times 1}$ at the downlink user is given by

$$\mathbf{y}_d = \mathbf{H}_d \mathbf{p} s_d + \mathbf{n}_d, \quad (1)$$

where $\mathbf{p} \in \mathbb{C}^{N_t \times 1}$ denotes the transmit precoder at the transceiver, $\mathbf{H}_d \in \mathbb{C}^{N_d \times N_t}$ the downlink communication channel, $\mathbf{n}_d \in \mathbb{C}^{N_d \times 1}$ the independent and identically distributed (i.i.d) additive complex Gaussian noise (i.e., $\mathbf{n}_d \sim \mathcal{CN}(\mathbf{0}, \sigma_d^2 \mathbf{I}_{N_d})$). At the downlink user, an estimate of s_d , denoted by \hat{s}_d , is obtained by filtering \mathbf{y}_d by a combiner $\mathbf{u}_d \in \mathbb{C}^{N_d \times 1}$, as follows:

$$\hat{s}_d = \mathbf{u}_d^H \mathbf{y}_d = \mathbf{u}_d^H \mathbf{H}_d \mathbf{p} s_d + \mathbf{u}_d^H \mathbf{n}_d. \quad (2)$$

At the FD ISAC transceiver, the received signal incorporates the SoI (i.e., the radar echo and the uplink symbol, s_u), along with residual SI. We assume that the received signal has no clipping error due to sufficient SI cancellation techniques in propagation and analog domains, and a high ADC dynamic range¹, of the order of 80dB, which can be achieved by a variety of ways, e.g., logarithmic ADC [37], modulo ADC [38], [39]). Hence, the received signal at the transceiver is given by

$$\mathbf{y}_u = \mathbf{H}_u \boldsymbol{\omega}_u s_u + \mathbf{H}_p s_d + \mathbf{n}_u, \quad (3)$$

where $\boldsymbol{\omega}_u \in \mathbb{C}^{N_u \times 1}$ denotes the precoder vector of the uplink user, $\mathbf{H}_u \in \mathbb{C}^{N_r \times N_u}$ the uplink communication channel,

¹The dynamic range is defined by the ratio between the largest and smallest possible values of the input signal, which are respectively the residual SI and radar echo in our system model. From our link budget simulation, the ratio between the residual SI and radar echo is 134.15dB, thus we assume that the SI cancellation techniques before digital domain can achieve larger than 54.15dB SI cancellation.

$\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ the aggregate channel comprising the radar channel, $\mathbf{H}_r \in \mathbb{C}^{N_r \times N_t}$, and the self-interference channel, $\mathbf{H}_{si} \in \mathbb{C}^{N_r \times N_t}$ (i.e., $\mathbf{H} = \mathbf{H}_r + \mathbf{H}_{si}$), and finally $\mathbf{n}_u \sim \mathcal{CN}(\mathbf{0}, \sigma_u^2 \mathbf{I}_{N_r}) \in \mathbb{C}^{N_r \times 1}$ the i.i.d additive complex Gaussian noise. The above-mentioned channels will be elaborated in Section II-B.

At the FD ISAC transceiver, an estimate of s_u , denoted by \hat{s}_u , is obtained by filtering \mathbf{y}_u by a receive beamformer $\mathbf{w} \in \mathbb{C}^{N_r \times 1}$ as follows:

$$\begin{aligned} \hat{s}_u &= \mathbf{w}^H \mathbf{y}_u \\ &= \underbrace{\mathbf{w}^H \mathbf{H}_u \boldsymbol{\omega}_u s_u}_{\text{Uplink stream}} + \underbrace{\mathbf{w}^H \mathbf{H}_r \mathbf{p} s_d}_{\text{Radar return}} + \underbrace{\mathbf{w}^H \mathbf{H}_{si} \mathbf{p} s_d}_{\text{Residual SI}} + \mathbf{w}^H \mathbf{n}_u, \end{aligned} \quad (4)$$

In Section III, we design beamformers \mathbf{w} and \mathbf{p} such that $\mathbf{w}^H \mathbf{H}_{si} \mathbf{p} \approx 0$ (i.e., SI suppression) while the uplink symbol and the radar returns are amplified. To decode s_u , we treat the interfering radar return as noise, while we use the favourable correlation properties of communications signals [40] to isolate the uplink data stream from the radar returns to estimate the target parameters. We provide more details regarding this in Section V.

Remark 1. *It is intuitive to choose a separate radar signal, and design two separate beamformers for radar and communications. However, for a fixed transmit power budget, this leads to the important question of power allocation across the communication and sensing symbols. Hence, it is not obvious that radar performance may be better with separate signals/precoders. Furthermore, for full-duplex operation, beamformer design with separate radar and communication signals is more complex as residual SI needs to be suppressed twice. For these reasons, we adopt the dual-functional transmit signal and a common beamformer.*

B. Channel Model

1) *Communication Channel:* The communication channels, \mathbf{H}_u and \mathbf{H}_d , are assumed to experience both small-scale and large-scale fading, which is modelled as follows, taking \mathbf{H}_u as an example

$$\mathbf{H}_u = \sqrt{\eta_u} \mathbf{G}_u, \quad (5)$$

where $\eta_u \in \mathbb{C}$ is the large-scale fading coefficient (including geometric attenuation and shadow fading), and $\mathbf{G}_u \in \mathbb{C}^{N_r \times N_u}$ is the small-scale fading matrix modelled by the classic Rician fading model [41], as follows

$$\mathbf{G}_u = \sqrt{\frac{\kappa}{\kappa+1}} \bar{\mathbf{G}}_u + \sqrt{\frac{1}{\kappa+1}} \tilde{\mathbf{G}}_u, \quad (6)$$

where κ (≥ 0) captures the proportion of the energy in the direct path (LoS), relative to the energy of the scattered paths (NLoS). We assume that the size of the antenna arrays is negligible compared with the distance between transceiver and user. Thus, the channel matrix corresponding to the LoS path (denoted by $\bar{\mathbf{G}}_u$ in (6)) is given by

$$\bar{\mathbf{G}}_u = \mathbf{a}(\theta_u) \mathbf{b}^\top(\theta'_u), \quad (7)$$

where

- θ_u and θ'_u denote the AoA at the transceiver and angle-of-departure (AoD) at the uplink user, respectively.
- $\mathbf{a}(\theta_u) = [1, e^{j2\pi\delta \sin(\theta_u)}, \dots, e^{j2\pi(N_r-1)\delta \sin(\theta_u)}]^\top \in \mathbb{C}^{N_r \times 1}$ is the steering vector in the direction of the AoA at the transceiver, with $\delta = 0.5$.
- $\mathbf{b}(\theta'_u) = [1, e^{j2\pi\delta \sin(\theta'_u)}, \dots, e^{j2\pi(N_u-1)\delta \sin(\theta'_u)}]^\top \in \mathbb{C}^{N_u \times 1}$ is the steering vector in the direction of the AoD at the uplink user.

Additionally, $\tilde{\mathbf{G}}_u$ is the i.i.d. complex Gaussian matrix, where each element follows $[\tilde{\mathbf{G}}_u]_{i,j} \sim \mathcal{CN}(0, 1)$.

2) *Radar Channel:* We assume that the radar operates in track model, where an initial estimate of the target's AoA is available from previous scanning. As in (3), the FD ISAC transceiver will receive a reflected echo with the target's AoA, range and velocity information (i.e., θ_r , r , v , respectively). Since we assume the transceiver to operate as a mono-static radar, the AoA and AoD are both equal to θ_r . We model the radar channel as varying slowly with time, with its expression given by

$$\mathbf{H}_r(t) = \eta_r e^{j2\pi f_d t} \mathbf{a}(\theta_r) \mathbf{b}^\top(\theta_r), \quad (8)$$

where $f_d = 2vf_c/c$ is the Doppler frequency, f_c and c are the signal carrier frequency and the speed of light. η_r is the attenuation factor, including round-trip path-loss and radar cross-section (RCS) σ_r , given in (9) derived from the radar range equation [42].

$$\eta_r = \sqrt{\frac{\lambda^2 \sigma_r}{(4\pi)^3 r^4}}. \quad (9)$$

3) *Self-Interference Channel:* In this work, we focus on the residual digital-domain SI cancellation assuming that the analog SI cancellation method is effective, i.e., there is no saturation at the ADC of the receive RF chains at the transceiver. To obtain \mathbf{H}_{si} , channel estimation methods based on classic estimation algorithms (e.g., least square (LS) [25], maximum-likelihood (ML) [43]) or machine learning-based methods [44] can be used. For illustration and without loss of generality, we consider a channel model derived from experiments [19], [45], which is similar to (6) and given by:

$$\mathbf{H}_{si} = \sqrt{\eta_{si}} \mathbf{G}_{si}, \quad (10)$$

where η_{si} and \mathbf{G}_{si} denote the large-scale fading coefficient and small-scale fading matrix, respectively. Specifically, \mathbf{G}_{si} has the same structure as \mathbf{G}_u in (6), containing LoS and NLoS components.

C. Problem Formulation

1) *Communication Rate:* With the uplink channel \mathbf{H}_u , radar channel \mathbf{H}_r , and SI channel \mathbf{H}_{si} in (5), (8), and (10), the uplink communication rate is given by

$$R_u = \log_2(1 + \gamma_u), \quad (11)$$

²For notation simplicity, we drop the time instant, (t) , in $\mathbf{H}_r(t)$.

where γ_u is the Signal-to-Interference-plus-Noise Ratio (SINR) of the received uplink signal. The expression for γ_u is given by

$$\gamma_u = \frac{|\mathbf{w}^H \mathbf{H}_u \boldsymbol{\omega}_u|^2}{|\mathbf{w}^H \mathbf{H}_p \mathbf{p}|^2 + \|\mathbf{w}\|^2 \sigma_u^2}. \quad (12)$$

Similarly, the downlink communication rate is given by

$$R_d = \log_2(1 + \gamma_d), \quad (13)$$

where γ_d is the SINR of the downlink signal given by

$$\gamma_d = \frac{|\mathbf{u}_d^H \mathbf{H}_d \mathbf{p}|^2}{\|\mathbf{u}_d\|^2 \sigma_d^2}. \quad (14)$$

In FD systems, in general, the downlink user would experience interference from uplink transmissions, as well. However, we assume that this interference is effectively managed by scheduling [46]–[48], and thus ignore its effect in (13).

2) *Radar Beampattern Power*: The transmit and receive beampattern power in the target direction, denoted by G_t and G_r , respectively, as a function of \mathbf{p} and \mathbf{w} , is given by

$$G_t = |\mathbf{b}^H(\theta_r) \mathbf{p}|^2, \quad G_r = |\mathbf{w}^H \mathbf{a}(\theta_r)|^2. \quad (15)$$

Remark 2. We use beampattern power as a proxy for radar SINR. While the latter may be a more intuitive sensing metric, its fractional form makes it non-convex – and thus, intractable – as well as difficult to obtain good bounds. On the other hand, the tractability of the beampattern power metric enables us to obtain closed-form solutions for key sub-problems (see Section III-B) that help Algorithm 1 converge faster.

3) *Objective Function and Constraints*: We aim to design the transmit and receive beamformer \mathbf{p} , \mathbf{w} at the transceiver, precoder $\boldsymbol{\omega}_u$ at the uplink user, and combiner \mathbf{u}_d at the downlink user to simultaneously maximize the uplink and downlink rate, the transmit and receive radar beampattern power at the target, and suppress the residual SI. In addition, the following constraints should be satisfied: 1) transmit power constraint at the transceiver and uplink user, i.e. $\|\mathbf{p}\|^2 \leq P_d$, $\|\boldsymbol{\omega}_u\|^2 \leq P_u$, where P_d and P_u are the maximum transmit power permitted at transceiver and uplink user, respectively; 2) receive power constraint at the transceiver, i.e. $\|\mathbf{w}\|^2 = 1$; 3) SI suppression constraint, i.e., $\mathbf{w}^H \mathbf{H}_{si} \mathbf{p} = 0$. With the above constraints, we can formulate the joint ISAC TX-RX beamformers design as

$$\max_{\mathbf{p}, \mathbf{w}, \boldsymbol{\omega}_u, \mathbf{u}_d} \alpha_1 R_u + \alpha_2 R_d + \alpha_3 G_t + \alpha_4 G_r \quad (16a)$$

$$\text{s.t.} \quad \|\mathbf{p}\|_2^2 \leq P_d, \quad (16b)$$

$$\|\boldsymbol{\omega}_u\|_2^2 \leq P_u, \quad (16c)$$

$$\|\mathbf{w}\|_2^2 = 1, \quad (16d)$$

$$\mathbf{w}^H \mathbf{H}_{si} \mathbf{p} = 0. \quad (16e)$$

The weight of the four components in (16a) is controlled by the coefficient $\alpha_n, \forall n \in \{1, 2, 3, 4\}$. The problem (16) is difficult to solve due to the following reasons: 1) the coupling between the transmit and receive beamformer in constraint (16e); 2) the non-convex communication metrics (i.e., $R_{u/d}$) in objective (16a); 3) the quadratic radar metrics (i.e. $G_{t/r}$) in objective (16a) in the maximization problem. To solve (16), we

transform it into an alternative form that permits an iterative solution with guaranteed convergence to a feasible solution.

III. JOINT TX-RX BEAMFORMER DESIGN

A. Problem Transformation

To address the difficulties in the previous section, we transform problem (16) based on the PDD method [49], the weighted Minimum Mean Square Error (MMSE) approach [36], and the reformulation of the quadratic beampattern power function [50]. Each of these techniques addresses a specific difficulty, as explained below:

1) Coupling Constraint Transformation by Penalty Term:

To tackle the coupling between \mathbf{w} and \mathbf{p} in constraint (16e), we transform the constraint $\mathbf{w}^H \mathbf{H}_{si} \mathbf{p} = 0$ into a penalty term $\frac{1}{2\beta} \|\mathbf{w}^H \mathbf{H}_{si} \mathbf{p}\|_2^2$, and add it to the objective function with the penalty parameter $\beta > 0$. In particular, β determines the penalty intensity of the digital-domain SI suppression, with lower values of β prioritizing SI suppression over communication and sensing performance. Thus, problem (16) can be reformulated as (17).

$$\begin{aligned} \max_{\{\mathbf{p}, \mathbf{w}\}, \{\boldsymbol{\omega}_u, \mathbf{u}_d\}} \quad & \alpha_1 R_u + \alpha_2 R_d + \alpha_3 G_t + \alpha_4 G_r \\ & - \frac{1}{2\beta} \|\mathbf{w}^H \mathbf{H}_{si} \mathbf{p}\|_2^2 \quad (17a) \\ \text{s.t.} \quad & (16b) - (16d). \end{aligned}$$

2) Communication Rate Transformation by WMMSE Method:

To circumvent the non-convexity of the rate terms in (17a), we adopt the weighted minimize mean square error (WMMSE) method to reformulate these terms based on the equivalence between the rate and MSE matrix [35], [36]. With the received signal (4) at the transceiver, the MSE matrix at the transceiver is defined as

$$\begin{aligned} E_{BS} &\triangleq \mathcal{E} \left\{ (\hat{\mathbf{s}}_u - \mathbf{s}_u) (\hat{\mathbf{s}}_u - \mathbf{s}_u)^H \right\} \\ &= \mathbf{w}^H \mathbf{H}_u \boldsymbol{\omega}_u \boldsymbol{\omega}_u^H \mathbf{H}_u^H \mathbf{w} - 2\Re\{\mathbf{w}^H \mathbf{H}_u \boldsymbol{\omega}_u\} \\ &\quad + \mathbf{w}^H (\mathbf{H}_p \mathbf{p} \mathbf{p}^H \mathbf{H}^H + \sigma_u^2 \mathbf{I}) \mathbf{w} + 1. \end{aligned} \quad (18)$$

Similarly, MSE matrix function of downlink user can be also defined as

$$\begin{aligned} E_d &\triangleq \mathcal{E} \left\{ (\hat{\mathbf{s}}_d - \mathbf{s}_d) (\hat{\mathbf{s}}_d - \mathbf{s}_d)^H \right\} \\ &= \mathbf{u}_d^H \mathbf{H}_d \mathbf{p} \mathbf{p}^H \mathbf{H}_d^H \mathbf{u}_d - 2\Re\{\mathbf{u}_d^H \mathbf{H}_d \mathbf{p}\} \\ &\quad + \sigma_d^2 \mathbf{u}_d^H \mathbf{u}_d + 1. \end{aligned} \quad (19)$$

Then, we introduce ρ_u and ρ_d as the auxiliary variables, and obtain the transformed objective function shown in (20) on the basis of rate-WMMSE relationship.

$$\begin{aligned} \max_{\substack{\{\mathbf{p}, \mathbf{w}\}, \{\boldsymbol{\omega}_u, \mathbf{u}_d\}, \\ \{\rho_u, \rho_d\}}} \quad & \alpha_1 (\log_2 \rho_u - \rho_u E_{BS}) + \alpha_2 (\log_2 \rho_d - \rho_d E_d) \\ & + \alpha_3 G_t + \alpha_4 G_r - \frac{1}{2\beta} \|\mathbf{w}^H \mathbf{H}_{si} \mathbf{p}\|_2^2 \quad (20a) \\ \text{s.t.} \quad & (16b) - (16d). \end{aligned}$$

3) *Beampattern Power Transformation*: To make G_t and G_r concave, we reformulate them with $V_t = \|\mathbf{b}(\theta_r)\|^2$ as

$$\begin{aligned} G_t &= |\mathbf{b}^H(\theta_r)\mathbf{p}|^2 = \mathbf{p}^H \mathbf{b}(\theta_r) \mathbf{b}^H(\theta_r) \mathbf{p} \\ &= \mathbf{p}^H \mathbf{b}(\theta_r) \mathbf{b}^H(\theta_r) \mathbf{p} - V_t P_d + V_t P_d \\ &\leq \mathbf{p}^H \mathbf{b}(\theta_r) \mathbf{b}^H(\theta_r) \mathbf{p} - \mathbf{p}^H V_t \mathbf{I} \mathbf{p} + V_t P_d \quad (21) \\ &= \mathbf{p}^H (\mathbf{b}(\theta_r) \mathbf{b}^H(\theta_r) - V_t \mathbf{I}) \mathbf{p} + V_t P_d \\ &= \mathbf{p}^H Z_t(\theta_r) \mathbf{p} + V_t P_d, \end{aligned}$$

where we denote $Z_t(\theta_r) = \mathbf{b}(\theta_r) \mathbf{b}^H(\theta_r) - V_t \mathbf{I}$. The equality holds when the power of \mathbf{p} reaches P_d . With the expression of the steering vector, it is straightforward to know that $\mathbf{b}(\theta_r) \mathbf{b}^H(\theta_r)$ is a rank-1 matrix with the eigenvalue $\|\mathbf{b}(\theta_r)\|^2$. Therefore, $Z_t(\theta_r)$ is negative semi-definite, and $\mathbf{p}^H Z_t(\theta_r) \mathbf{p}$ is concave. Subsequently, maximizing G_t can be approximated as the maximization of $\mathbf{p}^H Z_t(\theta_r) \mathbf{p}$ by ignoring the constant term $V_t P_d$.

Analogous to the reformulation of G_t , the expression of G_r is given by

$$G_r \leq \mathbf{w}^H Z_r(\theta_r) \mathbf{w} + V_r, \quad (22)$$

where $Z_r(\theta_r) = \mathbf{a}(\theta_r) \mathbf{a}_r^H(\theta_r) - V_r \mathbf{I}$ and $V_r = \|\mathbf{a}(\theta_r)\|^2$. The equality holds when $\|\mathbf{w}\|_2^2 = 1$. Consequently, problem (16) is reformulated as problem (23) shown at the top of the next page.

B. Penalty-based Block Update Algorithm

With the BCD method, an iterative approach is utilized where each variable is optimized while keeping the others fixed, till convergence.

1) *Block* $\{\rho_u, \rho_d\}$: When $\{\mathbf{p}, \mathbf{w}\}$, $\{\omega_u, \mathbf{u}_d\}$ are fixed, the two separate sub-problems with regards to ρ_u and ρ_d are both convex and unconstrained given by

$$\max_{\rho_u} \log_2 \rho_u - \rho_u E_{BS}, \quad (24)$$

and

$$\max_{\rho_d} \log_2 \rho_d - \rho_d E_d. \quad (25)$$

Thus, the optimal ρ_u^* and ρ_d^* can be obtained by setting the partial derivative with respect to the two optimized variables to zero. Subsequently, the expressions for ρ_u^* and ρ_d^* are:

$$\begin{aligned} \rho_u^* &= E_{BS}^{-1} \\ &= (\mathbf{w}^H \mathbf{H}_u \omega_u \omega_u^H \mathbf{H}_u^H \mathbf{w} - 2\Re\{\mathbf{w}^H \mathbf{H}_u \omega_u\} \\ &\quad + \mathbf{w}^H (\mathbf{H} \mathbf{p} \mathbf{p}^H \mathbf{H}^H + \sigma_u^2 \mathbf{I}) \mathbf{w} + 1)^{-1}, \end{aligned} \quad (26)$$

$$\begin{aligned} \rho_d^* &= E_d^{-1} \\ &= (\mathbf{u}_d^H \mathbf{H}_d \mathbf{p} \mathbf{p}^H \mathbf{H}_d^H \mathbf{u}_d - 2\Re\{\mathbf{u}_d^H \mathbf{H}_d \mathbf{p}\} \\ &\quad + \sigma_d^2 \mathbf{u}_d^H \mathbf{u}_d + 1)^{-1}. \end{aligned} \quad (27)$$

The complexity of (26) and (27) are $\mathcal{O}(N_r^2)$, and $\mathcal{O}(N_t N_d)$, respectively, mainly due to matrix multiplication.

2) *Block* $\{\omega_u\}$: With fixed $\{\mathbf{p}, \mathbf{w}\}$, $\{\rho_u, \rho_d\}$, the sub-problem with regard to ω_u is given by

$$\begin{aligned} \min_{\omega_u} \quad & \alpha_1 E_{BS} \\ \text{s.t.} \quad & \|\omega_u\|_2^2 \leq P_u. \end{aligned} \quad (28)$$

Since the objective function is convex with the convex constraint, the Lagrange multiplier method based on the Karush–Kuhn–Tucker (KKT) condition is used with expression:

$$\min_{\omega_u} \alpha_1 E_{BS} + \mu (\|\omega_u\|_2^2 - P_u). \quad (29)$$

To solve this problem, we first substitute (29) with (18). Subsequently, we take the partial derivative of the Lagrangian function with regard to ω_u and μ , respectively, and set both to zero. As a result, we obtain

$$\omega_u = (\alpha_1 \mathbf{w}^H \mathbf{H}_u \mathbf{H}_u^H \mathbf{w} + \mu)^{-1} \alpha_1 \mathbf{H}_u^H \mathbf{w}, \quad (30)$$

and

$$\|\omega_u\|_2^2 = P_u. \quad (31)$$

By substituting (31) with (30), we can then utilize the bisection search to obtain μ^* . Thereon, the optimal solution ω_u^* is obtained by replacing μ with μ^* in (30).

$$\omega_u^* = (\alpha_1 \mathbf{w}^H \mathbf{H}_u \mathbf{H}_u^H \mathbf{w} + \mu^*)^{-1} \alpha_1 \mathbf{H}_u^H \mathbf{w}. \quad (32)$$

Updating ω_u requires $\mathcal{O}(I_1 N_r N_u)$ complexity due to the bisection search for μ with I_1 the number of iterations and the matrix multiplication.

3) *Block* $\{\mathbf{w}\}$: Under the condition that $\mathbf{p}, \omega_u, \mathbf{u}_d, \rho_u$, and ρ_d are all fixed, \mathbf{w} is only related to the first and the last two terms in (23a), yielding the following convex sub-problem

$$\min_{\mathbf{w}} \alpha_1 \rho_u E_{BS} + \alpha_4 \mathbf{w}^H Z_r(\theta_r) \alpha_4 \mathbf{w} + \frac{1}{2\beta} \|\mathbf{w}^H \mathbf{H}_{si} \mathbf{p}\|_2^2. \quad (33)$$

Although the constraint (16d) that $\|\mathbf{w}\|_2^2 = 1$ is a non-convex constraint, we firstly address this sub-problem with respect to \mathbf{w} as an unconstrained complex optimization problem, and normalize the obtained solution after the convergence of the algorithm is satisfied. Equating the partial derivative of (33) with respect to \mathbf{w} to zero, the optimal \mathbf{w}^* is given by

$$\mathbf{w}^* = (2\mathbf{X}_u + \frac{1}{\beta} \mathbf{H}_{si} \mathbf{p} \mathbf{p}^H \mathbf{H}_{si}^H)^{-1} 2\alpha_1 \rho_u \mathbf{H}_u \omega_u, \quad (34)$$

where \mathbf{X}_u is defined as

$$\begin{aligned} \mathbf{X}_u &= \alpha_1 \rho_u \mathbf{H}_u \omega_u \omega_u^H \mathbf{H}_u^H + \alpha_1 \rho_u \mathbf{H} \mathbf{p} \mathbf{p}^H \mathbf{H}^H \\ &\quad + \alpha_1 \rho_u \sigma_u^2 \mathbf{I} - \alpha_4 Z_r(\theta_r). \end{aligned} \quad (35)$$

The complexity of this step is $\mathcal{O}(N_r^3)$ mainly due to the matrix inversion.

4) *Block* $\{\mathbf{u}_d\}$: The uplink precoder is only related to the second term in (23a) with other blocks fixed. Therefore, the unconstrained convex sub-problem is given by

$$\begin{aligned} \min_{\mathbf{u}_d} \quad & \mathbf{u}_d^H \mathbf{H}_d \mathbf{p} \mathbf{p}^H \mathbf{H}_d^H \mathbf{u}_d - 2\Re\{\mathbf{u}_d^H \mathbf{H}_d \mathbf{p}\} \\ & + \sigma_d^2 \mathbf{u}_d^H \mathbf{u}_d + 1. \end{aligned} \quad (36)$$

The optimal solution is given by

$$\mathbf{u}_d^* = (\mathbf{H}_d \mathbf{p} \mathbf{p}^H \mathbf{H}_d^H + \sigma_d^2 \mathbf{I})^{-1} \mathbf{H}_d \mathbf{p}. \quad (37)$$

Updating the \mathbf{u}_d in (37) requires $\mathcal{O}(N_t N_d)$ complexity mainly due to matrix multiplication.

$$\begin{aligned} & \max_{\{\mathbf{p}, \mathbf{w}\}, \{\boldsymbol{\omega}_u, \mathbf{u}_d\}, \{\rho_u, \rho_d\}} \alpha_1(\log_2 \rho_u - \rho_u \mathbf{E}_{BS}) + \alpha_2(\log_2 \rho_d - \rho_d \mathbf{E}_d) + \alpha_3 \mathbf{p}^H Z_t(\theta_t) \mathbf{p} + \alpha_4 \mathbf{w}^H Z_r(\theta_r) \mathbf{w} - \frac{1}{2\beta} \|\mathbf{w}^H \mathbf{H}_{si} \mathbf{p}\|_2^2 \quad (23a) \\ & \text{s.t.} \quad (16b) - (16d). \end{aligned}$$

5) *Block* $\{\mathbf{p}\}$: Under fixed $\rho_u, \rho_d, \mathbf{w}, \boldsymbol{\omega}_u$ and \mathbf{u}_d , the sub-problem with respect to \mathbf{p} is a convex optimization problem with a convex constraint on the transceiver transmit power, P_d , as follows

$$\begin{aligned} \min_{\mathbf{p}} \quad & \alpha_1 \rho_u \mathbf{E}_{BS} + \alpha_2 \rho_d \mathbf{E}_d - \alpha_3 \mathbf{p}^H Z_t(\theta_t) \mathbf{p} + \frac{1}{2\beta} \|\mathbf{w}^H \mathbf{H}_{si} \mathbf{p}\|_2^2 \\ \text{s.t.} \quad & \|\mathbf{p}\|_2^2 \leq P_d. \end{aligned} \quad (38)$$

Similar to the design of block $\{\boldsymbol{\omega}_u\}$ in III-B2, this sub-problem with respect to \mathbf{p} can be reformulated by the Lagrange multiplier method based on the KKT condition written below:

$$\begin{aligned} \min_{\mathbf{p}} \quad & \alpha_1 \rho_u \mathbf{E}_{BS} + \alpha_2 \rho_d \mathbf{E}_d - \alpha_3 \mathbf{p}^H Z_t(\theta_t) \mathbf{p} + \frac{1}{2\beta} \|\mathbf{w}^H \mathbf{H}_{si} \mathbf{p}\|_2^2 \\ & + \Gamma (\|\mathbf{p}\|_2^2 - P_d). \end{aligned} \quad (39)$$

By taking the partial derivative with respect to Γ and \mathbf{p} , respectively, and setting them to zero, we obtain

$$\|\mathbf{p}\|_2^2 = P_d, \quad (40)$$

and

$$\begin{aligned} \mathbf{p} = & (2\mathbf{X}_d + \frac{1}{\beta} (\mathbf{H}_{si}^H \mathbf{w} \mathbf{w}^H \mathbf{H}_{si}) + 2\Gamma \mathbf{I})^{-1} \\ & \cdot (2\alpha_2 \rho_d \mathbf{H}_d^H \mathbf{u}_d), \end{aligned} \quad (41)$$

where \mathbf{X}_d is defined in (42) for expression simplicity,

$$\mathbf{X}_d = \alpha_1 \rho_u \mathbf{H}^H \mathbf{w} \mathbf{w}^H \mathbf{H} + \alpha_2 \rho_d \mathbf{H}_d \mathbf{u}_d \mathbf{u}_d^H \mathbf{H}_d^H - \alpha_3 Z_t(\theta_t) \mathbf{I}. \quad (42)$$

Subsequently, we substitute (41) into (40), and utilize the bisection search to obtain the optimal Γ^* . Finally, the optimal \mathbf{p}^* is given by

$$\begin{aligned} \mathbf{p}^* = & (2\mathbf{X}_d + \frac{1}{\beta} (\mathbf{H}_{si}^H \mathbf{w} \mathbf{w}^H \mathbf{H}_{si}) + 2\Gamma^* \mathbf{I})^{-1} \\ & \cdot (2\alpha_2 \rho_d \mathbf{H}_d^H \mathbf{u}_d). \end{aligned} \quad (43)$$

The complexity of updating \mathbf{p} in (43) is $\mathcal{O}(I_2 N_t^3)$ caused by the matrix inversion, where I_2 denotes the number of iterations in the bisection search for Γ .

6) *Summary*: Following the framework of penalty-based approach, iterative BCD update process, and setting appropriate initial value of the optimized variables, $\rho_u, \rho_d, \mathbf{w}, \boldsymbol{\omega}_u, \mathbf{u}_d$, and \mathbf{p} , the optimal solution in each block can be found after reaching the convergence condition. To satisfy the receive power constraint of \mathbf{w} , we normalize \mathbf{w} after the whole iteration as below:

$$\mathbf{w} = \frac{\mathbf{w}}{\|\mathbf{w}\|}. \quad (44)$$

We summarize our proposed joint ISAC TX-RX beamformers design algorithm in Algorithm 1.

Algorithm 1: Proposed Penalty-based Joint Transmit and Receive Beamformer Design

Input: $\mathbf{H}_d, \mathbf{H}_u, \mathbf{H}_r, \mathbf{H}_{si}, P_d, P_u$.

Output: $\rho_u^*, \rho_d^*, \mathbf{p}^*, \mathbf{w}^*, \boldsymbol{\omega}_u^*, \mathbf{u}_d^*$.

- 1 Initialize $\mathbf{p}, \mathbf{w}, \boldsymbol{\omega}_u, \mathbf{u}_d$ randomly, $\beta = 10^{-25}$;
 - 2 **while** no convergence of objective function (23a) **do**
 - 3 Update ρ_u^* by (26). Complexity = $\mathcal{O}(N_r^2)$.
 - 4 Update ρ_d^* by (27). Complexity = $\mathcal{O}(N_t N_d)$.
 - 5 Update $\boldsymbol{\omega}_u^*$ by (32). Complexity = $\mathcal{O}(I_1 N_r N_u)$.
 - 6 Update \mathbf{w}^* by (34). Complexity = $\mathcal{O}(N_r^3)$.
 - 7 Update \mathbf{u}_d^* by (37). Complexity = $\mathcal{O}(N_t N_d)$.
 - 8 Update \mathbf{p}^* by (43). Complexity = $\mathcal{O}(I_2 N_t^3)$.
 - 9 **end**
 - 10 Normalize \mathbf{w}^* by (44);
 - 11 **Return** $\rho_u^*, \rho_d^*, \mathbf{p}^*, \mathbf{w}^*, \boldsymbol{\omega}_u^*, \mathbf{u}_d^*$.
-

Remark 3. An alternative to (16e) is to restrict the residual SI power to below a threshold (i.e., $\frac{1}{2\beta} \|\mathbf{w}^H \mathbf{H}_{si} \mathbf{p}\|_2^2 < \varepsilon$), which has been shown to improve communications performance in [51]. However, we have adopted a stronger constraint in (16e) because with the inclusion of sensing, where the radar returns are both very weak and strongly correlated with the residual SI, the latter needs to be suppressed to a far greater extent.

Nevertheless, enforcing such a strong constraint deteriorates both communications and sensing performance. Our compromise to this trade-off is detailed in the following comment.

Remark 4. Comparison between PDD and Algorithm 1 PDD is a widely-used technique to solve optimization problems involving non-smooth, non-convex functions [49]. Algorithm 1 is similar to PDD due to the same problem formulation as shown in (17a). However, the PDD is a two-loop iterative algorithm, wherein the inner loop solves an augmented Lagrangian problem by BCD method (i.e., the steps within the while loop in Algorithm 1), and the outer loop updates the dual variable and penalty term until convergence (i.e., $\lambda \mathbf{w}^H \mathbf{H}_{si} \mathbf{p} + \frac{1}{2\beta} \|\mathbf{w}^H \mathbf{H}_{si} \mathbf{p}\|_2^2 < \varepsilon$). In Algorithm 1, we omit the outer loop iteration of the PDD framework because for sufficiently small β , the residual SI power can be suppressed below a suitable threshold (e.g. noise floor). In contrast, having the outer loop would needlessly suppress the residual SI further (i.e., enforcing (16e)), which not only reduces both communications and sensing performance, but also leads to slower convergence.

Remark 5. In general, the ISAC TX-RX beamformer design problem can be formulated in several ways; for instance, an alternative to (16a)-(16e) is to maximize communications

performance to constraints on radar sensing performance.

However, an advantage of our formulation is that it permits a closed-form solution for each step of the BCD (i.e., each of the subproblems in Section III-B). On the other hand, incorporating the radar sensing performance as a constraint would have made it difficult to have realize this, and thus slowed down the convergence of our Algorithm. On the other hand, a softer constraint on sensing performance can be imposed by increasing the weights α_3 and α_4 , which would have the effect of prioritizing the sensing performance.

Remark 6. The joint TX-RX beamformer design can be extended to underloaded multi-user scenarios as well. For example, consider K downlink users and let \mathbf{p}_k denote the downlink precoder for the k^{th} downlink user and $\mathbf{P} := [\mathbf{p}_1, \dots, \mathbf{p}_K] \in \mathbb{C}^{N_t \times K}$. For $K < \min(N_t, N_r)$, the rows of $\mathbf{H}_{\text{si}}\mathbf{P}$ have a non-trivial nullspace and hence, there exists non-zero \mathbf{w} for which $\mathbf{w}^H \mathbf{H}_{\text{si}}\mathbf{P} = 0$.

C. Algorithm Implementation

In this paper, we consider a centralized implementation of Algorithm 1, which involves the overhead of communicating all the channel state information ($\mathbf{H}_u, \mathbf{H}_{\text{si}}, \mathbf{H}_d$) and the target direction to a controller as well as transferring the optimized \mathbf{p} and \mathbf{w} to the ISAC transceiver, ω_u to the uplink user and ω_d to the downlink user. Thus, a decentralized implementation of Algorithm 1 is key for minimizing the overhead. Furthermore, the order in which the variables are optimized can be different, and thus, a different ordering along with possibly different update rates may speed up convergence. These questions are left for future work. In the following section, we discuss the convergence behavior and the complexity of Algorithm 1.

IV. CONVERGENCE AND COMPLEXITY ANALYSIS

A. Convergence Analysis

In this section, the convergence of Algorithm 1 to at least a local optimum is proved. Let $f(\mathbf{p}, \mathbf{w}, \omega_u, \mathbf{u}_d)$ denote the objective function in (17a), and $f_\rho(\mathbf{p}, \mathbf{w}, \omega_u, \mathbf{u}_d, \rho_u, \rho_d)$ the reformulated objective function in (20a), based on the rate-WMMSE relationship explained in Section III-A2. Thus, we have

$$f(\mathbf{p}, \mathbf{w}, \omega_u, \mathbf{u}_d) \geq f_\rho(\mathbf{p}, \mathbf{w}, \omega_u, \mathbf{u}_d, \rho_u, \rho_d), \quad (45)$$

where ρ_u^* , and ρ_d^* are the optimal solutions as shown in (26) and (27). The inequality condition in (45) holds when ρ_u satisfies (26) and ρ_d satisfies (27) are both the global optimum with fixed other optimized variables $\mathbf{p}, \mathbf{w}, \omega_u, \mathbf{u}_d$. Let $f_g(\mathbf{p}, \mathbf{w}, \omega_u, \mathbf{u}_d, \rho_u, \rho_d)$ denote the objective in (23a). Then we have

$$f_\rho(\mathbf{p}, \mathbf{w}, \omega_u, \mathbf{u}_d, \rho_u, \rho_d) \leq f_g(\mathbf{p}, \mathbf{w}, \omega_u, \mathbf{u}_d, \rho_u, \rho_d) + \alpha_3 V_t P_d + \alpha_4 V_r. \quad (46)$$

Let $\mathbf{p}^{(n)}, \mathbf{w}^{(n)}, \omega_u^{(n)}, \mathbf{u}_d^{(n)}, \rho_u^{(n)}, \rho_d^{(n)}$ be the n -th iteration of the variables whose expressions are given by (43), (34), (32), (37), (26) and (27), respectively. $\rho_u^{(n)}, \rho_d^{(n)}$ are obtained by the n -th iteration beamformers $\{\mathbf{p}^{(n)}, \mathbf{w}^{(n)}, \omega_u^{(n)}\}$ and

$\{\mathbf{p}^{(n)}, \mathbf{u}_d^{(n)}\}$, respectively. Since the power of $\mathbf{p}^{(n)}$ reaches P_d , we obtain (47) from (45) and (46), given by

$$\begin{aligned} & f(\mathbf{p}^{(n+1)}, \mathbf{w}^{(n+1)}, \omega_u^{(n+1)}, \mathbf{u}_d^{(n+1)}) \\ & \geq f_g(\mathbf{p}^{(n+1)}, \mathbf{w}^{(n+1)}, \omega_u^{(n+1)}, \mathbf{u}_d^{(n+1)}, \rho_u^{(n)}, \rho_d^{(n)}) \\ & \quad + \alpha_3 V_t P_d + \alpha_4 V_r. \end{aligned} \quad (47)$$

Following Algorithm 1, the update sequences are $\omega_u^{(n+1)} \stackrel{(a)}{\leftarrow} \{\omega_u^{(n)}\}$, $\mathbf{w}^{(n+1)} \stackrel{(b)}{\leftarrow} \{\mathbf{p}^{(n)}, \omega_u^{(n)}, \rho_u^{(n)}\}$, $\mathbf{u}_d^{(n+1)} \stackrel{(c)}{\leftarrow} \{\mathbf{p}^{(n)}\}$, and $\mathbf{p}^{(n+1)} \stackrel{(d)}{\leftarrow} \{\rho_u^{(n)}, \rho_d^{(n)}, \mathbf{w}^{(n+1)}, \mathbf{u}_d^{(n+1)}\}$. Therefore, we have

$$\begin{aligned} & f_g(\mathbf{p}^{(n+1)}, \mathbf{w}^{(n+1)}, \omega_u^{(n+1)}, \mathbf{u}_d^{(n+1)}, \rho_u^{(n)}, \rho_d^{(n)}) \\ & \geq f_g(\mathbf{p}^{(n)}, \mathbf{w}^{(n)}, \omega_u^{(n)}, \mathbf{u}_d^{(n)}, \rho_u^{(n)}, \rho_d^{(n)}). \end{aligned} \quad (48)$$

Subsequently, by using (45) and (46) and due to the power of $\mathbf{p}^{(n)}$ reaching P_d , we have

$$\begin{aligned} & f_g(\mathbf{p}^{(n)}, \mathbf{w}^{(n)}, \omega_u^{(n)}, \mathbf{u}_d^{(n)}, \rho_u^{(n)}, \rho_d^{(n)}) + \alpha_3 V_t P_d + \alpha_4 V_r \\ & = f(\mathbf{p}^{(n)}, \mathbf{w}^{(n)}, \omega_u^{(n)}, \mathbf{u}_d^{(n)}). \end{aligned} \quad (49)$$

From (45) - (49), we obtain

$$f(\mathbf{p}^{(n+1)}, \mathbf{w}^{(n+1)}, \omega_u^{(n+1)}, \mathbf{u}_d^{(n+1)}) \geq f(\mathbf{p}^{(n)}, \mathbf{w}^{(n)}, \omega_u^{(n)}, \mathbf{u}_d^{(n)}), \quad (50)$$

which shows that the objective function in (17a) is non-decreasing after each BCD update iteration. Since the uplink and downlink rates are upper bounded [52], as are the beam-pattern gains when beamformers \mathbf{p} and \mathbf{w} both probe at target direction, θ_r , the value of objective in (17a) is bounded above. Consequently, the proposed algorithm will converge to at least a local optimum, which guarantees the convergence of the proposed Algorithm 1. In the following, we provide numerical results to demonstrate the convergence of the proposed design.

Fig. 2 shows the behavior of the objective function (17a) and the relative difference ζ , defined in (51), with respect to the number of iterations, for different N_t and SI power levels.

$$\zeta \triangleq \frac{f_g^{(n)} - f_g^{(n-1)}}{f_g^{(n-1)}}, \quad (51)$$

where $\zeta \leq \epsilon$ is the iteration end criterion. As shown in Fig. 2, objective (23a) converges within a small number of iterations.

B. Complexity Analysis

From Algorithm 1, the bottleneck step is the step 8, which updates \mathbf{p} , with assumption $N_d < N_t \leq N_r$, $N_u < N_t \leq N_r$. Thus, the overall complexity of Algorithm 1 is given by

$$\mathcal{O}(I_3(N_r^2 + N_t N_d + I_1 N_r N_u + N_r^3 + I_2 N_t^3)) \approx \mathcal{O}(I_3(I_2 N_t^3)), \quad (52)$$

where I_3 is the iteration number. In summary, Algorithm 1 has polynomial complexity.

V. NUMERICAL EVALUATION

In this section, we provide numerical results to validate the performance of the proposed joint TX-RX beamformer design for the FD ISAC system. The number of transmit and receive antennas at the FD ISAC transceiver are set the same, which are $N_t = N_r = 16$. The number of antennas at

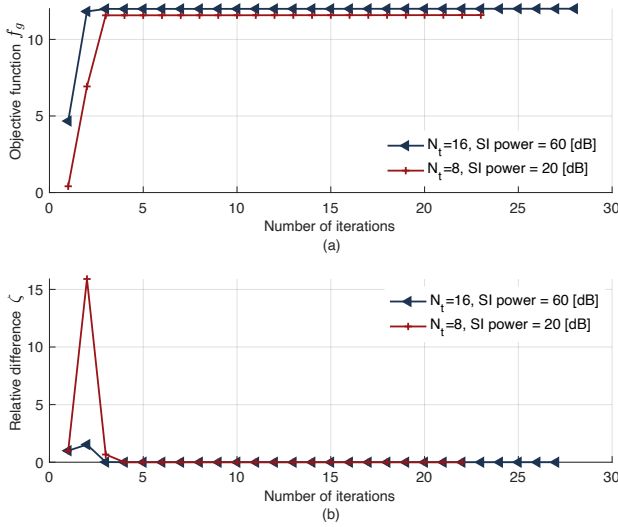


Fig. 2. Convergence of Algorithm 1 with $\epsilon = 10^{-5}$. (a) Objective function versus the number of iterations; (b) Relative difference versus the number of iterations.

uplink and downlink user are set to $N_u = 2$ and $N_d = 2$, respectively. The carrier frequency of the FD ISAC system is $f_c = 2.4\text{GHz}$. The sampling rate is $T_s = 1/\Omega$, where $\Omega = 20\text{MHz}$ is the bandwidth. We use the tuple (θ, r, v) to represent a target/user's coordinates and bearing w.r.t the transceiver, where θ denotes the direction, r the range and v the velocity. The radar target is assumed to be at $(45^\circ, 7.5\text{m}, 20\text{m/s})$, the uplink user at $(-50^\circ, 10\text{m}, 0\text{m/s})$, and the downlink user at $(\theta_d = -30^\circ, 100\text{m}, 0\text{m/s})$. The transmit powers of the ISAC transceiver and uplink user are $P_{d,[\text{dBm}]} = 20\text{dBm}$, and $P_{u,[\text{dBm}]} = 10\text{dBm}$, respectively. The thermal noise floor at the ISAC receiver is $P_{\text{noise},[\text{dBm}]} = -94\text{dBm}$. The path loss (in dB) is modelled as follows $\eta(d) = -20 \log_{10}(\lambda/(4\pi d_0)) + 10n \log_{10}(d/d_0)$, where $d_0 = 1\text{m}$ is the reference distance, and $n = 2.2$ is the path loss exponent [53], respectively. The SI power, $P_{\text{si},[\text{dBm}]} - P_{\text{noise},[\text{dBm}]}$, ranges from 10 - 60dB at the ISAC receiver. We run 1000 times Monte Carlo simulations with perfect channel state information (CSI) for observing the performances of the SI cancellation, SoI over SI, and sum-rates. The signals, s_u , and s_d , comprise uncoded i.i.d zero mean, unit-energy QPSK symbols.

A. SI Power After Cancellation

Firstly, we measure the SI cancellation performance of Algorithm 1 in terms of the residual digital-domain SI power level $P_{\text{res},[\text{dB}]}$, which is defined as $P_{\text{res},[\text{dB}]} := 10 \log_{10}(|\mathbf{w}^H \mathbf{H}_{\text{si}} \mathbf{p}|^2 / 1\text{mW}) - P_{\text{noise},[\text{dBm}]}$. Thus, a negative value of P_{res} implies that the residual SI power level is lower than the noise floor. Additionally, we assume $\alpha_{\text{com}} = \alpha_1 = \alpha_2$ (i.e., equal weightage for uplink and downlink communications performance) and $\alpha_{\text{radar}} = \alpha_3 = \alpha_4$. Furthermore, the priority given to communications performance is captured by the parameter, $\varrho = \alpha_{\text{com}}/\alpha_{\text{radar}}$. As shown in Fig. 3, when the ϱ varies from 1 to 1000, the SI power up to 60 dB with regards to noise floor can be effectively suppressed due to the residual SI power, $P_{\text{res},[\text{dB}]} < 0$. To observe whether the SoI

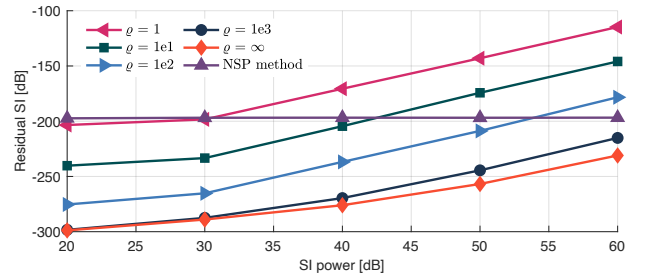


Fig. 3. Performance comparison of SI cancellation with varying priority parameter ϱ and level of SI. The residual SI level is computed with respect to the noise floor.

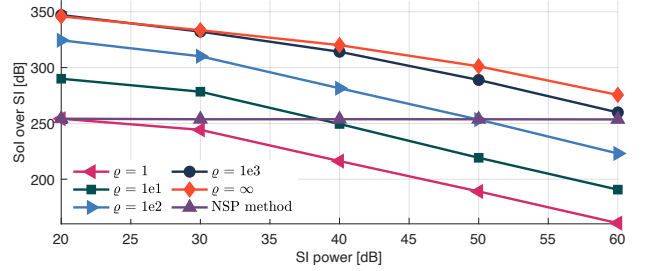


Fig. 4. Performance comparison of SoI (e.g., radar echo and uplink data) power over SI power in dB.

is preserved in the process of SI suppression, we consider the ratio between the SoI and SI powers, defined as $P_{\text{SoI}/\text{SI},[\text{dB}]} := 10 \log_{10}((|\mathbf{w}^H \mathbf{H}_r \mathbf{p}|^2 + |\mathbf{w}^H \mathbf{H}_u \omega_u|^2) / |\mathbf{w}^H \mathbf{H}_{\text{si}} \mathbf{p}|^2)$. The performance of $P_{\text{SoI}/\text{SI},[\text{dB}]}$ with varying ϱ is shown in Fig. 4, where a positive value implies that the SoI (e.g., radar echo and uplink data) is not drowned by the strong SI. As expected, the performance of SI cancellation degrades and $P_{\text{SoI}/\text{SI},[\text{dB}]}$ decreases with increasing of the SI power. Additionally, when the priority parameter ϱ increases, the performance of SI cancellation improves and SoI over SI power increases. This is because a higher ϱ leads to more priority given to communications, resulting in reduced correlation between SI and SoI. It should be noted that this correlation is one of the challenges in suppressing SI when the sensing function is present. Thus, the joint ISAC TX-RX beamformers design can effectively preserve SoI and suppress SI in FD ISAC. While NSP method has better performances than the proposed method with some certain parameter settings, other system capabilities of the proposed method (i.e., average sum-rate of downlink and uplink users, radar parameter estimation performance) are enhanced, which are explained in Section V-C and Section V-D, respectively.

B. Beam pattern Power Performance

Fig 5 illustrate examples of the transmit and receive beam patterns with varying priority parameters ϱ with effective cancellation of 60dB residual SI. As shown in Figs. 5a and 5d, when $\varrho = 1$, the transmit beamformer focuses the transmit power towards the radar target and the downlink user directions. Meanwhile, the receive beamformer concentrates on the uplink user and radar target. According to Figs. 5b

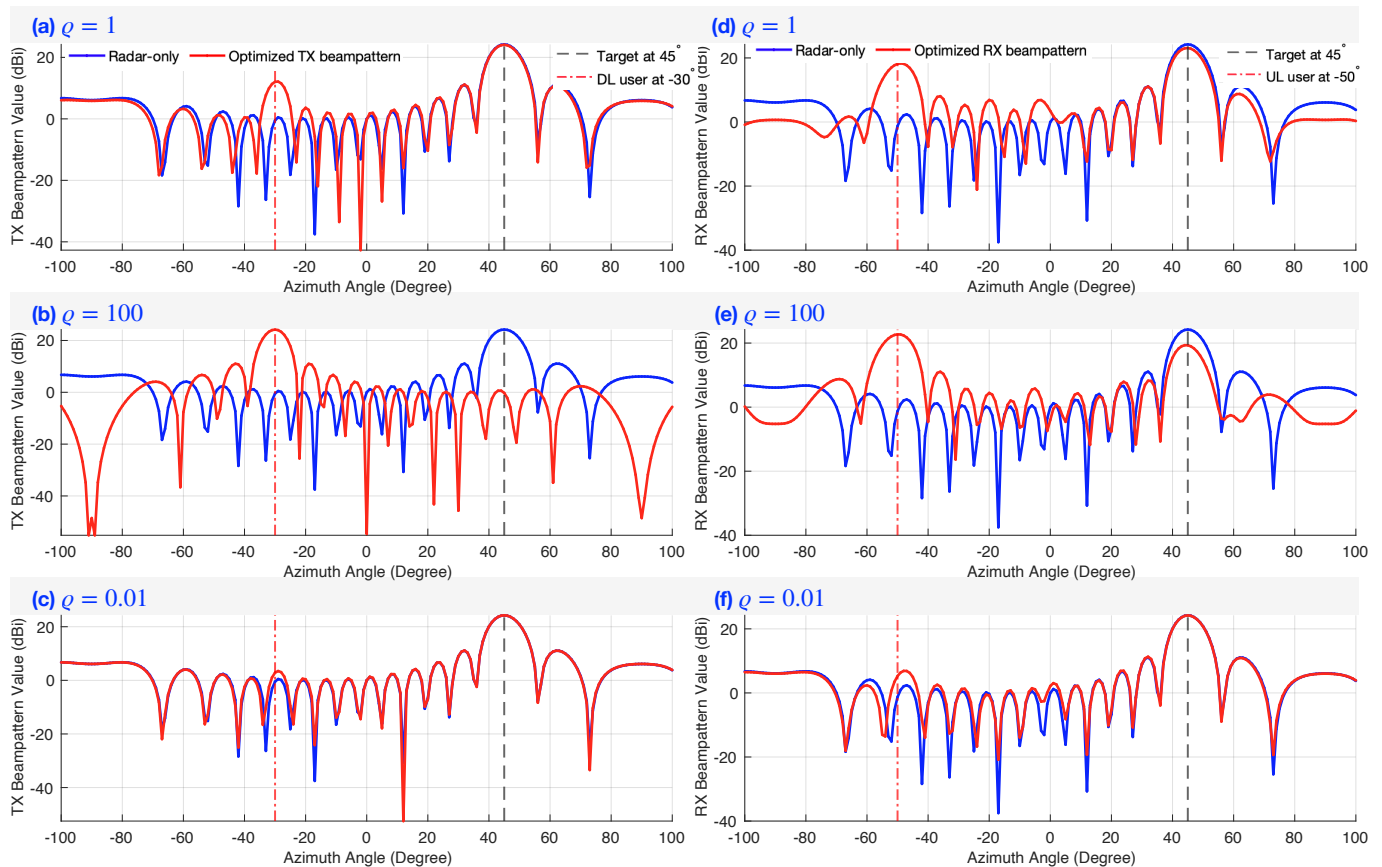


Fig. 5. **Left panel:** Transmit beampattern, **Right panel:** Receive beampattern of the ISAC FD system with an effective cancellation of 60 dB SI, where $N_t = 16$, target at $\theta_r = 45^\circ$, and downlink user at $\theta_d = -30^\circ$: (a, d) $\rho = 1$; (b, e) $\rho = 100$; (c, f) $\rho = 0.01$.



Fig. 6. Frame structures of (a) half-duplex and (b) FD systems.

and 5e, the transmit and receive beams focus more on the downlink user and uplink user compared with the transmit and receive beampatterns of $\rho = 1$, respectively, due to a higher priority on the communication that is $\rho = 100$. When $\rho = 0.01$, the radar function has higher priority, thus the transmit and receive beams concentrate more on the target in comparison with the transmit and receive beampatterns of $\rho = 1$. Additionally, the residual SI power level $P_{\text{res, [dB]}}$ for the cases that $\rho = 1, 100, 0.01$ are -130.97dB , -197.37dB , and -108.48dB , respectively. Hence, the SI is effectively suppressed.

C. Sum-rate Performance

In half-duplex systems, a fraction, δ , of the (time-frequency) resources are allocated for uplink and the remaining for downlink, while for FD, all the resources are allocated simultaneously for both uplink and downlink, as shown in Fig. 6.

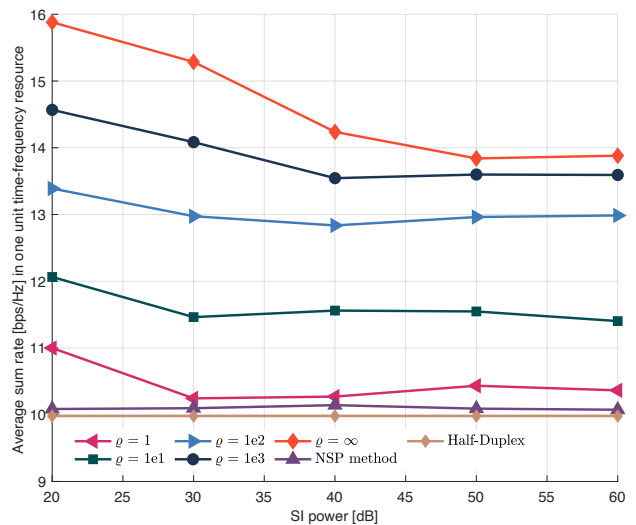


Fig. 7. Average sum-rate in one unit time-frequency resource versus varying level of SI power.

The sum-rate in one unit time-frequency resource for the half-duplex system is defined as $R_{\text{hd, [unit]}} = \delta R_{\text{dl}} + (1 - \delta) R_{\text{ul}}$, and sum-rate for the FD system is defined as $R_{\text{fd, [unit]}} = R_{\text{dl}} + R_{\text{ul}}$. In our simulation, δ is set as 0.5, which means equal time-

frequency resource allocation for both uplink and downlink.

The sum-rate performance with respect to SI power level with varying priority parameter ϱ is shown in Fig. 7. When ϱ increases, the sum-rate increases correspondingly. Specifically, the upper bound is given by the sum-rate of the FD communication functions (i.e., $\varrho = \infty$). The lower bound is given by the sum-rate in one unit time-frequency resource of the half-duplex system. With increasing SI power, the sum-rate in unit time-frequency resource first decreases and then tends to be flat thanks to the effective suppression of the residual SI.

D. Sensing Performance

For range-Doppler sensing, we explicitly consider the signal stream in (4) in discrete time, ignoring the residual SI. The resulting signal can then be represented as

$$\begin{aligned} \hat{s}_u[n, m] &= \mathbf{w}^H \mathbf{H}_u \boldsymbol{\omega}_u s_u[n, m] \\ &+ \eta_r e^{j2\pi(f_d T_s)m} \mathbf{w}^H \mathbf{a}(\theta_r) \mathbf{b}^\top(\theta_r) \mathbf{p} s_d[n - i_\tau, m] \\ &+ \mathbf{w}^H \mathbf{n}_u[n, m], \\ &(n = 0, \dots, N; m = 0, \dots, M - 1), \end{aligned} \quad (53)$$

where the indices m and n respectively capture the slow and fast time-scales commonly assumed in range-Doppler processing³, and $i_\tau \in \mathbb{Z}$ is the round-trip delay of the radar echo (also known as the target range bin). The signal stream, $s_d[n, m] := s_d[n + mM]$ (and likewise, with $s_u[n, m]$) can be viewed as a concatenation of M blocks, with each block comprising N symbols.

The range bin, i_τ , is estimated from (53) by applying the matched filter w.r.t $s_d[n, m]$ along the fast-time axis. In [40], it was shown that the cross-correlation function between $s_u[\cdot, m]$ and $s_d[\cdot, m]$, and the autocorrelation function of $s_d[\cdot, m]$ asymptotically converged to the all-zero function and $\delta[\cdot]$, respectively, for large N . Hence, the uplink data stream has negligible impact on the range-Doppler sensing performance. Subsequently, we perform an M -point DFT operation, resulting in the range-Doppler map.

With respect to angular domain sensing, since we implicitly assume that the target is in the vicinity of the direction in which \mathbf{w} is "pointing" towards (i.e., the radar operates in *track mode*), we are more interested in the interference suppression capability of the latter from an unwanted direction, θ , captured by the following metric:

$$P(\theta) = |\mathbf{w}^H (\mathbf{a}(\theta) + \mathbf{H}_{\text{si}} \mathbf{p})|^2. \quad (54)$$

Consequently, the AoA estimate is given by $\arg \max_{\theta} P(\theta)$.

For comparison, we consider the NSP method, radar-only method, and communication-only method. Specifically, the TX and RX beamformers for the radar-only method are given by $\mathbf{p}_r = \mathbf{b}(\theta_r)$, and $\mathbf{w}_r = \mathbf{a}(\theta_r)$, respectively. Similarly, the TX and RX beamformers for the communication-only method are given by $\mathbf{p}_u = \mathbf{b}(\theta_u)$, and $\mathbf{w}_u = \mathbf{a}(\theta_u)$, respectively. The priority parameter ϱ is set to 1. The range-velocity map is shown

³This framework results in a decoupling between the effects of the delay and Doppler shifts in (53), which is a reasonable assumption when the Doppler frequency (320Hz in our case) is much smaller than the signal bandwidth (20MHz in our case).

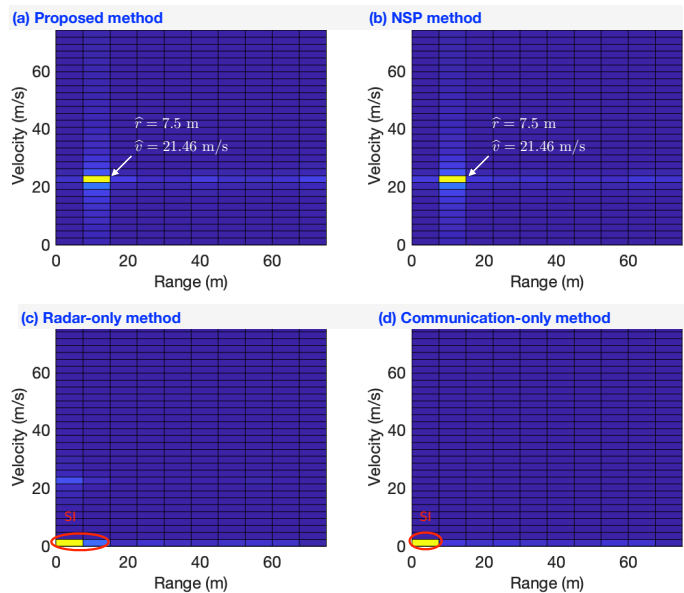


Fig. 8. Examples of range and velocity estimation based on the FD ISAC system with 60 dB SI with $N_t = 16$, $N_r = 16$, $M = 512$ blocks, and $N = 1024$ symbols.

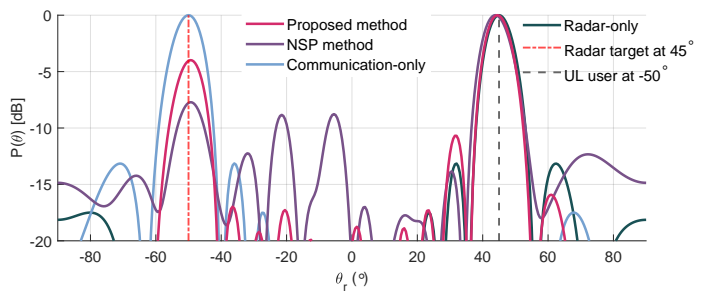


Fig. 9. An example of angle estimation for a simulated scenario with a target at $\theta_r = 45^\circ$, a uplink communication user at $\theta_u = -50^\circ$ with antennas $N_t = 16$ and 60 dB SI power.

in Fig. 8. We observe that Algorithm 1 has effective residual SI suppression compared with methods that do not consider SI suppression (i.e., the radar-only and communications-only methods), where the desired radar echo drowns in the SI, which can be seen as a strong signal with $v = 0$ m/s, $r = 0$ m. Additionally, the radar-only and communication-only methods both have interference around the original point due to the unsuppressed residual SI.

Fig. 9 contains plots of the output power, $P(\theta)$, of different methods. We observe that the proposed method has a close angle detection output power, $P(\theta)$, at the target direction but a distinct output power at the uplink user direction, when compared with the radar-only method. The proposed method also has a similar output power compared with communication-only method. Additionally, the proposed method has fewer interference compared with the NSP method, especially around the angle 0° , where the SI locates.

VI. CONCLUSION

In this paper, we design the transmit and receive beamformer \mathbf{p} , \mathbf{w} at the transceiver, precoder ω_u at the uplink user, combiner \mathbf{u}_d at the downlink user to simultaneously maximize the uplink and downlink rate, the transmit and receive radar beampattern power at the target, and suppress the residual SI. In the objective function, TX and RX beampattern gains are used as the radar metric, and the uplink and downlink rates are used as the communication metric. With the aid of the equivalence of the rate maximization and the MSE minimization, and penalty-based transformation, we use the BCD method to solve the optimization problem. Subsequently, we give a convergence analysis. Numerical results show that up to 60 dB residual SI in digital domain can be efficiently suppressed. Additionally, the optimized TX and RX beampatterns can probe at the desired target, uplink and downlink user directions with a satisfactory average sum-rate, a more accurate radar parameter estimation with regards to range, velocity, and angle, and outperforms the NSP beamformer design method, which validates the effectiveness of the proposed algorithm. For future studies, a more effective algorithm on suppressing SI is worthy investigating. In addition, an extended transceiver design for a more general multi-data transmission can be studied. Moreover, it would be interesting to extend the SI cancellation technique in digital domain to full domain in view of more accurate residual SI model and saturation caused by analog SI.

REFERENCES

- [1] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, 2019.
- [2] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.
- [3] N. Su, F. Liu, and C. Masouros, "Secure radar-communication systems with malicious targets: Integrating radar, communications and jamming functionalities," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 83–95, 2020.
- [4] M. F. Keskin, V. Koivunen, and H. Wymeersch, "Limited feedforward waveform design for OFDM dual-functional radar-communications," *IEEE Trans. Signal Process.*, vol. 69, pp. 2955–2970, 2021.
- [5] K. Wu, J. A. Zhang, X. Huang, and Y. J. Guo, "Integrating low-complexity and flexible sensing into communication systems," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1873–1889, 2022.
- [6] X. Liu, H. Zhang, K. Long, M. Zhou, Y. Li, and H. V. Poor, "Proximal policy optimization-based transmit beamforming and phase-shift design in an IRS-aided ISAC system for the THz band," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2056–2069, 2022.
- [7] Q. Shi, L. Liu, S. Zhang, and S. Cui, "Device-free sensing in OFDM cellular network," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1838–1853, 2022.
- [8] X. Lin, J. Li, R. Baldemair, J.-F. T. Cheng, S. Parkvall, D. C. Larsson, H. Koorapaty, M. Frenne, S. Falahati, A. Grovlen *et al.*, "5G new radio: Unveiling the essentials of the next generation wireless access technology," *IEEE Commun. Stand. Mag.*, vol. 3, no. 3, pp. 30–37, 2019.
- [9] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, 2014.
- [10] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, 2014.
- [11] Z. Zhang, K. Long, A. V. Vasilakos, and L. Hanzo, "Full-duplex wireless communications: Challenges, solutions, and future research directions," *Proc. IEEE*, vol. 104, no. 7, pp. 1369–1409, 2016.
- [12] Z. Xiao and Y. Zeng, "Waveform design and performance analysis for full-duplex integrated sensing and communication," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1823–1837, 2022.
- [13] Y. He, Y. Cai, G. Yu, and K.-K. Wong, "Joint transceiver design for dual-functional full-duplex relay aided radar-communication systems," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8355–8369, 2022.
- [14] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," *Proc. IEEE*, vol. 99, no. 7, pp. 1236–1259, 2011.
- [15] P. Kumari, S. A. Vorobyov, and R. W. Heath, "Adaptive virtual waveform design for millimeter-wave joint communication–radar," *IEEE Trans. Signal Process.*, vol. 68, pp. 715–730, 2019.
- [16] P. Kumari, J. Choi, N. González-Prelcic, and R. W. Heath, "IEEE 802.11 ad-based radar: An approach to joint vehicular communication–radar system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3012–3027, 2017.
- [17] C. B. Barneto, T. Riihonen, M. Turunen, L. Anttila, M. Fleischer, K. Stadius, J. Ryyänen, and M. Valkama, "Full-duplex OFDM radar with LTE and 5G NR waveforms: Challenges, solutions, and measurements," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 10, pp. 4042–4054, 2019.
- [18] K. E. Kolodziej, B. T. Perry, and J. S. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3025–3041, 2019.
- [19] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, 2012.
- [20] K. E. Kolodziej, J. G. McMichael, and B. T. Perry, "Multitap RF canceller for in-band full-duplex wireless communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4321–4334, 2016.
- [21] E. Everett, C. Shepard, L. Zhong, and A. Sabharwal, "SoftNull: Many-antenna full-duplex wireless via digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8077–8092, 2016.
- [22] B. Debaillie, D.-J. van den Broek, C. Lavin, B. van Liempd, E. A. Klumperink, C. Palacios, J. Craninckx, B. Nauta, and A. Pärssinen, "Analog/RF solutions enabling compact full-duplex radios," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1662–1673, 2014.
- [23] D. Korpi, J. Tamminen, M. Turunen, T. Huusari, Y.-S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Full-duplex mobile device: Pushing the limits," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 80–87, 2016.
- [24] S. Hong, J. Brand, J. I. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, "Applications of self-interference cancellation in 5G and beyond," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 114–121, 2014.
- [25] E. Ahmed and A. M. Eltawil, "All-digital self-interference cancellation technique for full-duplex systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3519–3532, 2015.
- [26] K. Komatsu, Y. Miyaji, and H. Uehara, "Basis function selection of frequency-domain hammerstein self-interference canceller for in-band full-duplex wireless communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3768–3780, 2018.
- [27] I. Hwang, B. Song, C. Nguyen, and S. S. Soliman, "Digitally controlled analog wideband interference cancellation for in-device spectrum sharing and aggregation," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2838–2850, 2016.
- [28] C. B. Barneto, S. D. Liyanaarachchi, T. Riihonen, M. Heino, L. Anttila, and M. Valkama, "Beamforming and waveform optimization for OFDM-based joint communications and sensing at mm-Waves," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2020, pp. 895–899.
- [29] S. D. Liyanaarachchi, C. B. Barneto, T. Riihonen, M. Heino, and M. Valkama, "Joint multi-user communication and MIMO radar through full-duplex hybrid beamforming," in *2021 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S)*. IEEE, 2021, pp. 1–5.
- [30] C. B. Barneto, T. Riihonen, S. D. Liyanaarachchi, M. Heino, N. González-Prelcic, and M. Valkama, "Beamformer design and optimization for full-duplex joint communication and sensing at mm-Waves," *arXiv preprint arXiv:2109.05932*, 2021.
- [31] P. K. Sharma and P. Garg, "Intelligent reflecting surfaces to achieve the full-duplex wireless communication," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 622–626, 2020.
- [32] A. M. T. Khel and K. A. Hamdi, "Performance analysis of ired-assisted full-duplex wireless communication systems with interference," *IEEE Commun. Lett.*, vol. 26, no. 9, pp. 2027–2031, 2022.
- [33] J. M. B. da Silva, A. Sabharwal, G. Fodor, and C. Fischione, "1-bit phase shifters for large-antenna full-duplex mmWave communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6916–6931, 2020.
- [34] M.-M. Zhao, Q. Wu, M.-J. Zhao, and R. Zhang, "Exploiting amplitude control in intelligent reflecting surface aided wireless communication

- with imperfect CSI," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4216–4231, 2021.
- [35] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, 2008.
- [36] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, 2011.
- [37] J. Lee, J. Kang, S. Park, J. S. Seo, J. Anders, J. Guilherme, and M. P. Flynn, "A 2.5 mW 80 dB DR 36 dB SNDR 22 MS/s logarithmic pipeline ADC," *IEEE J. Solid-State Circuits*, vol. 44, no. 10, pp. 2755–2765, 2009.
- [38] Z. Liu, A. Bhandari, and B. Clerckx, " λ -MIMO: Massive MIMO via modulo sampling," *arXiv preprint arXiv:2210.10193*, 2022.
- [39] A. Bhandari, F. Kraemer, and R. Raskar, "On unlimited sampling and reconstruction," *IEEE Trans. Signal Process.*, 2020.
- [40] S. Aditya, O. Dizdar, B. Clerckx, and X. Li, "Sensing using coded communications signals," *arXiv preprint arXiv:2209.04269*, 2022.
- [41] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
- [42] M. I. Skolnik, "Introduction to radar systems," *New York*, 1980.
- [43] A. Masmoudi and T. Le-Ngoc, "A maximum-likelihood channel estimator for self-interference cancellation in full-duplex systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5122–5132, 2015.
- [44] K. Muranov, M. A. Islam, B. Smida, and N. Devroye, "On deep learning assisted self-interference estimation in a full-duplex relay link," *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2762–2766, 2021.
- [45] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-aho, "On the spectral efficiency of full-duplex small cell wireless systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4896–4910, 2014.
- [46] J. Bai and A. Sabharwal, "Distributed full-duplex via wireless side-channels: Bounds and protocols," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4162–4173, 2013.
- [47] A. Sahai, S. Diggavi, and A. Sabharwal, "On degrees-of-freedom of full-duplex uplink/downlink channel," in *2013 IEEE Information Theory Workshop (ITW)*. IEEE, 2013, pp. 1–5.
- [48] C. Karakus and S. Diggavi, "Opportunistic scheduling for full-duplex uplink-downlink networks," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 1019–1023.
- [49] Q. Shi and M. Hong, "Penalty dual decomposition method for non-smooth nonconvex optimization—Part I: Algorithms and convergence analysis," *IEEE Trans. Signal Process.*, vol. 68, pp. 4108–4122, 2020.
- [50] C. Xu, B. Clerckx, and J. Zhang, "Multi-antenna joint radar and communications: Precoder optimization and weighted sum-rate vs probing power tradeoff," *IEEE Access*, vol. 8, pp. 173 974–173 982, 2020.
- [51] O. Taghizadeh, J. Zhang, and M. Haardt, "Transmit beamforming aided amplify-and-forward MIMO full-duplex relaying with limited dynamic range," *Signal Processing*, vol. 127, pp. 266–281, 2016.
- [52] B. Clerckx and C. Oestges, *MIMO wireless networks: channels, techniques and standards for multi-antenna, multi-user and multi-cell systems*. Academic Press, 2013.
- [53] H. Xu, V. Kukshya, and T. S. Rappaport, "Spatial and temporal characteristics of 60-GHz indoor channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 620–630, 2002.