Sensing-Assisted Channel Estimation in ISAC Systems

Rui Yin[1(✉)], Yexin Shi[2], Wei Qi[1], Xianfu Chen[3], Celimuge Wu[4], and Yusheng Ji[5]

[1] School of Information and Electrical Engineering, Hangzhou City University, Hangzhou 310018, China
yinrui@zucc.edu.cn

[2] College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China
2210600@zju.edu.cn

[3] Shenzhen CyberAray Network Technology Company Ltd, Shenzhen, China
xianfu.chen@ieee.org

[4] Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo 182-8585, Japan
celimuge@uec.ac.jp

[5] Information Systems Architecture Research Division, National Institute of Informatics, Tokyo 101-8430, Japan
kei@nii.ac.jp

Abstract. In future Sixth-Generation (6G) mobile communication, satellite communication is a crucial means of extending mobile networks worldwide. However, in satellite communication systems, the transmission capacity between ground stations and users has become a bottleneck for satellite communication systems. For this issue, Integrated Sensing and Communication (ISAC) technology and Artificial Intelligence (AI), which has been widely studied, has the potential to offer excellent solutions. This paper explores a Uformer-based sensing-assisted communication system that enhances channel estimation accuracy through the design of a Uformer-based channel estimation enhancement algorithm. By formulating a global optimization problem, we seek to improve resource efficiency through finding the optimal resource allocation schemes. Initially, the system utilizes radar data to obtain reconstructed channels and employs communication data for channel estimation. The proposed neural network model then fuses these data into a new and more accurate channel state information. Finally, we present a resource allocation algorithm and validate the performance of all proposed algorithms through simulations.

Keywords: Integrated sensing and communication · OFDM-MIMO · Artificial intelligence · Channel estimation

H.-H. Chen and W. Meng (Eds.): WiSATS 2024, LNICST 606, pp. 74–88, 2025.
https://doi.org/10.1007/978-3-031-86203-8_7

1 Introduction

In the planning for the Sixth-Generation (6G) wireless communications, satellite communication is viewed as a solution offering broad coverage, high reliability, and diverse services [1]. With its extensive coverage and low-latency communication, satellite communication can support network connections in remote areas, filling the gaps left by terrestrial networks and ensuring seamless global connectivity [2].

However, in common satellite communication schemes, the cost for users to directly access satellites is high, and in the scheme where user's signals are relayed to satellites via Base Stations (BS), the communication quality between the user and the ground BS becomes a bottleneck [3]. In the future development of 6G, Integrated Sensing And Communication (ISAC) technology is considered an effective method to improve communication service quality between BS and users. ISAC aims to integrate the hardware and signal processing process of sensing and communication systems to reduce system deployment costs and achieve improvement for both systems [4]. In communication-centric ISAC systems, the high-precision positioning, mapping, and environmental monitoring capabilities provided by the sensing system can offer rich environmental information and assist in various traditional communication tasks, such as beamforming and Channel Estimation (CE) [5,6].

Some research focuses on how to use sensing data to enhance the accuracy of CE in communication [7]. For instance, [8] investigated sensing-assisted CE methods by sparse recovery. In [8], the delay, Doppler shift, and angle estimation of communication channel paths were inferred by collecting radar signal echoes. The Orthogonal Matching Pursuit (OMP) algorithm was then used to recover the true channel from the initial CE results. Additionally, [9] employed a Kalman filtering method to estimate Channel State Information (CSI) based on angle estimation from radar data. These studies attempt to obtain sensing data, such as delays and angles of all signal paths, and then reconstruct the channel based on the multi-path channel model. However, these traditional methods only considered the main paths detected by radar, neglecting the weaker Non-Line-Of-Sight (NLOS) paths that are also part of the original CSI. Therefore, [10] discussed the deep learning method in joint CE and feedback problem.

In this paper, we focus on the channel estimation problem and propose an Uformer-based Channel Fusion Algorithm (UCFA). This algorithm can generate more accuracy channel than radar reconstruction channel and estimated CSI by fusing them. The neural network called Uformer is migrated from the field of image processing to CE in our study [11]. On this basis, we propose a new system framework and formulate two resource allocation problems for uplink and downlink to maximize the transmission rate.

The remainder of this paper is organized as follows: In Sect. 2, we introduce our proposed system, including the signal model and the system workflow. Then, a Uformer-based channel fusion algorithm is detailed in Sect. 3. Subsequently, Sect. 4 analyzes the performance metrics of the system and establish optimization problems for both uplink and downlink process and the solutions are proposed

in Sect. 5. In Sect. 6, the performance of the proposed UCFA and optimization algorithms is validated through numerical simulations. Finally, Sect. 7 concludes the paper.

It is worth mentioning that we use the following notations: $\mathbf{A}^{\mathrm{H}}$, $\mathbf{A}^{-1}$, $\mathrm{tr}(\mathbf{A})$, $\det(\mathbf{A})$ denote the Hermitian, inverse, trace, determinant of matrix $\mathbf{A}$.

2 System Model

In this section, we discuss the communication and transmission between ground BS and users in satellite communication. Any satellite communication link must rely on a ground station to achieve sufficiently strong transmission power and receiving capabilities. However, the communication between the ground station and the user is similar to traditional wireless networks, and the transmission performance will be affected by the environment. Therefore, we aim to study methods to improve the accuracy of channel estimation.

Taking a roadside BS as an example, the cell covered by this BS includes K users. Communication between the users and the BS employs Orthogonal Frequency Division Multiplexing (OFDM) technology, where each user is assigned a subchannel orthogonal to those of other users, thus avoiding mutual interference. Additionally, each user is equipped with a Uniform Linear Array (ULA) consisting of N_U antennas, which alternately perform sensing and transmission tasks in time-division protocol, with transmission including both uplink and downlink. Similarly, the BS is equipped with N_B antennas for communication with the users. Since the BS can typically afford higher deployment costs, we assume that the equipments for satellite communication, user communication, and environmental sensing at the BS operates independently. This design enhances the BS's sensing capabilities, enabling it to undertake critical tasks such as traffic management, emergency response, and crowd monitoring.

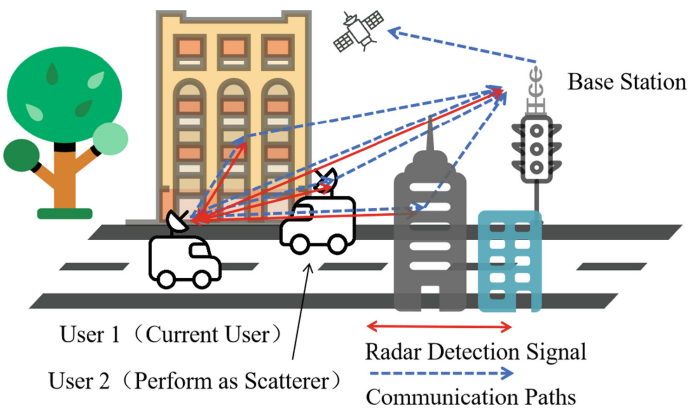

Fig. 1. The ISAC system model in satellite communication

As shown in Fig. 1, after the user transmits sensing signals, it can receive and analyze the echo signals to obtain parameter estimates of the scatterers, including other users and the buildings. Based on the delay and angle, triangulation methods can then be used to determine the relative positions of the various scatterers in the environment with respect to the user. In this work, we assume that the antennas of both the user and the BS are at the same height, allowing them to be modeled on a two-dimensional plane. When the user communicates with the BS, the signal transmitted by one side can reach the other side through various paths, which can be classified into three types: Line-of-Sight (LOS) paths, single-hop NLOS paths, and multi-hop NLOS paths. The LOS paths are direct paths that reach the target without any reflections, while single-hop and multi-hop paths are distinguished by the number of reflections the signal undergoes.

2.1 Signa Model

In this section, we will model the sensing signals and communication signals separately, taking the uplink as example. Since both rely on the environment, they are highly coupled.

Sensing Part. Ignoring the possible false alarming and miss detection, all M scatterers are assumed to be detected by one UE. Labeling all the scatterers by the numbers in the set $\mathcal{M} = \{1, 2, .., M\}$, and the echo received by the user can be expressed as:

$$\mathbf{Y}_S = \underbrace{\alpha_n \mathbf{a}_{\mathrm{R}}(\phi_{\mathrm{T},n})\mathbf{a}_{\mathrm{T}}^{\mathrm{H}}(\phi_{\mathrm{T},n})\mathbf{X}_S}_{\text{Target}} + \underbrace{\sum_{m=0,m\neq n}^{M} \alpha_m \mathbf{a}_{\mathrm{R}}(\phi_{\mathrm{T},m})\mathbf{a}_{\mathrm{T}}^{\mathrm{H}}(\phi_{\mathrm{T},m})\mathbf{X}_S}_{\text{Clutter}} + \mathbf{Z}_S, \tag{1}$$

where the received echo signal, $\mathbf{Y}_S$, is determined by the transmitting signal $\mathbf{X}_S = \mathbf{F}_S\mathbf{S}_S \in \mathbb{C}^{N_{\mathrm{U}}\times 1}$, while the transmitted signal vector, $\mathbf{S}_S$, is shaped by the beamforming matrix $\mathbf{F}_S$. In the sensing channel, α_m denotes the path gain coefficient of the m-th path, and $\mathbf{Z}_S$ represents the additive white Gaussian noise with variance σ_S^2. Here, we label the reflection paths generated by scatterers with the same indices as those of scatterers, while 0 are used to denote the LOS path. Additionally, $\mathbf{a}_{\mathrm{R}}(\phi)$ and $\mathbf{a}_{\mathrm{T}}(\phi)$ are the array steering vectors at the receiver and transmitter, respectively:

$$\mathbf{a}_{\mathrm{T}}(\phi) = \frac{1}{\sqrt{N_{\mathrm{T}}}}[\exp\{j(-\frac{N_{\mathrm{T}}-1}{2})\frac{2\pi d}{\lambda}\sin(\phi)\}, ..., \exp\{j(\frac{N_{\mathrm{T}}-1}{2})\frac{2\pi d}{\lambda}\sin(\phi)\}]^T, \tag{2}$$

$$\mathbf{a}_{\mathrm{R}}(\phi) = \frac{1}{\sqrt{N_{\mathrm{R}}}}[\exp\{j(-\frac{N_{\mathrm{R}}-1}{2})\frac{2\pi d}{\lambda}\sin(\phi)\}, ..., \exp\{j(\frac{N_{\mathrm{R}}-1}{2})\frac{2\pi d}{\lambda}\sin(\phi)\}]^T, \tag{3}$$

where N_{R} and N_{T} denote the antenna numbers of receiver and transmitter. In the case of a mono-static radar, the transmitter and receiver are the same and thus the two numbers are both equal to $N_{\mathbf{U}}$. Because the echo returns along the incident path, the Angle of Arrival (AOA) and Angle of Departure (AOD) are identical. We denote the angle of the m-th path as $\phi_{\mathrm{T},m}$. Moreover, d represents the antenna spacing, and λ denotes the wavelength.

Communication Part. During the communication process, all scatterers can generate single-hop paths, which we also label with the same identifiers as the scatterers themselves. In fact, some of these paths may be obstructed or experience significant path loss, nearly disappearing. To address this, we can set the corresponding path gains to 0 or small values. Additionally, we assume the existence of R multi-hop paths, which are also labeled with numbers from the set $\mathcal{R} = \{M+1, ..., M+R\}$. Thus, the received signal can be expresses as:

$$\begin{aligned}\mathbf{Y}_C &= \underbrace{\beta_0 \mathbf{b}_{\mathrm{R}}(\phi_{\mathrm{R},0})\mathbf{a}_{\mathrm{T}}^{\mathrm{H}}(\phi_{\mathrm{T},0})\mathbf{F}_C\mathbf{S}_C}_{\text{LOS}} + \underbrace{\sum_{m\in\mathcal{R}\cup\mathcal{P}} \beta_m \mathbf{b}_{\mathrm{R}}(\phi_{\mathrm{R},m})\mathbf{a}_{\mathrm{T}}^{\mathrm{H}}(\phi_{\mathrm{T},m})\mathbf{F}_C\mathbf{S}_C}_{\text{NLOS}} + \mathbf{Z}_C \\ &= \mathbf{H}_C\mathbf{F}_C\mathbf{S}_C + \mathbf{Z}_C, \end{aligned} \tag{4}$$

where $\mathbf{Y}_C, \mathbf{H}_C, \mathbf{F}_C, \mathbf{S}_C, \mathbf{Z}_C$ are similar to $\mathbf{Y}_S, \mathbf{H}_S, \mathbf{F}_S, \mathbf{S}_S, \mathbf{Z}_S$, distinguished only by the subscripts S and C to indicate whether they belong to the sensing or communication system. Besides, β_m represents the channel gain coefficient of m-th path in communication. Unlike the sensing part, where the signals are reflected back to the same device, in the communication process, the signals eventually arrive at a different device, the ground BS. Consequently, the scatter path differs from the incident path, and the Angle of Arrival (AOA) for the m-th path is denoted as $\phi_{\mathbf{R},m}$ and the corresponding receiving steering vector can be written as:

$$\mathbf{b}_{\mathrm{R}}(\phi) = \frac{1}{\sqrt{N_{\mathrm{R}}}}[\exp\{j(-\frac{N_{\mathrm{R}}-1}{2})\frac{2\pi d}{\lambda}\sin(\phi)\}, ..., \exp\{j(\frac{N_{\mathrm{R}}-1}{2})\frac{2\pi d}{\lambda}\sin(\phi)\}]^T. \tag{5}$$

In this context, N_{R} corresponds to the number of antennas at the BS, denoted as N_B. For both our analysis and simulations, we assume that all antenna spacings are identical.

2.2 Framework of the System

In traditional time-division duplexing (TDD) communication, uplink and downlink transmissions occur sequentially, with antennas switching between transmission and reception. However, in ISAC systems, users allocate time within each transmission frame cycle to perform sensing tasks. First, users reconstruct the channel based on sensing data to obtain sensing CSI (sCSI). During transmission, the base station performs channel estimation to obtain communication

CSI (cCSI), which is then fed back to the user via downlink transmission. The user inputs both sCSI and cCSI into a neural network, and the algorithm produces an output, fCSI. This fCSI is more accurate than sCSI and cCSI obtained through traditional methods and can be used to guide beamforming, timing, and power allocation schemes in the next transmission frame. The specific process is illustrated in Fig. 2.

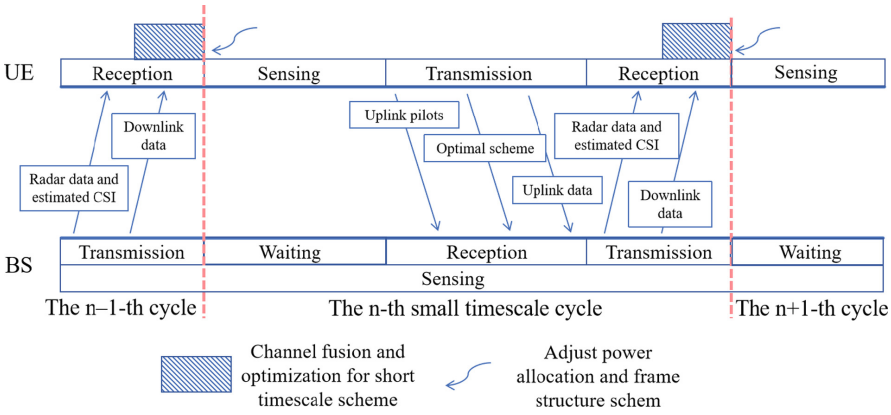

Fig. 2. The framework of TDD cycle.

During the transmission process of the entire cycle, all symbols can be classified into three categories: uplink pilot symbols, uplink data symbols, and downlink data symbols. In downlink transmission, the BS does not send pilots to estimate the channel but instead obtains the downlink channel through channel reciprocity. When the user receive the cCSI feedback and BS's radar data, the UCFA and resource allocation algorithms can work and then the generated optimal scheme can be applied in the next cycle.

During the sensing process, the system can obtain the localization of all scatterers and the BS, and then can calculate the angle of arrival, angle of departure. Besides, the channel gain of each path can be recovered from the cCSI by OMP or Multiple Signal Classification (MUSIC) algorithm. Finally, based on the multi-path model and all the parameters, the sCSI can be reconstructed.

3 Uformer-Based Channel Fusion Algorithm

The original Uformer network is a model designed for image denoising and deblurring based on a U-shaped network structure that incorporates the self-attention mechanism of the Transformer. In the context of channel estimation, when both the user and the base station are equipped with linear arrays, the CSI is a two-dimensional complex matrix, which can also be viewed as a two-channel two-dimensional image. Therefore, this paper aims to use Uformer as the basic

structure and design peripheral components for channel fusion tasks. The main structure of our proposed algorithm is illustrated in the Fig. 3.

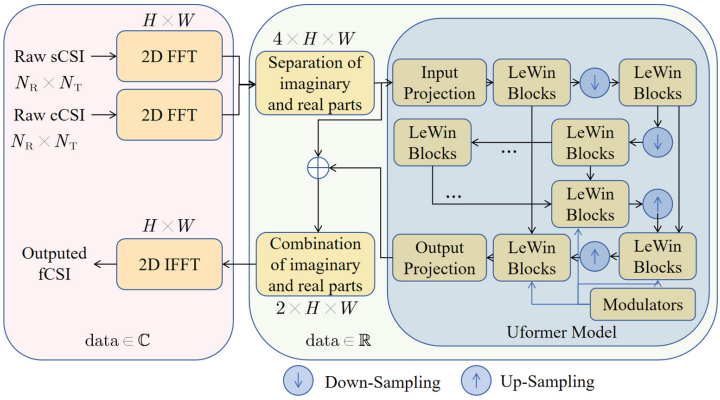

Fig. 3. The structure of the Uformer-based algorithm.

As shown in the Fig. 3, the raw cCSI and sCSI will be send to the data pre-processing parts which consists of three components: 2D Fourier transform, amplitude normalization, and real-imaginary separation. Firstly, the raw CSI lacks distinct features, but the 2D Fourier transform can capture the spatial spectral characteristics of the channel. This step can enhance the image details and enlarge the matrix size. For image processing models, the input matrix size is usually several hundred pixels, while the number of antennas for mobile users is at most a few dozen. Therefore, enlarging the matrix size helps fit the network. Additionally, the Fourier transform itself does not affect the noise power, so the denoising results of the transformed matrix can be fully reflected in the original matrix. Then, the amplitude normalization can improve the generalization performance of the network. And finally, the data will be separated into real and imaginary parts, then reorganized into a four-channel real matrix format before being input into the pre-process part. When the main part output the result which is the residual matrix, the algorithm should combine it with the input and the inverse process of pre-processing takes place.

4 Performance Analysis and Problem Formulation

Since we will use the fCSI generated by the neural network to design the beamforming matrix, how to ensure that the neural network algorithm can meet the channel estimation accuracy requirements becomes a key issue. Obviously, if too few resources are allocated to sensing or transmitting pilots, the accuracy of cCSI and sCSI will drop significantly, thus limiting the performance of fCSI. Therefore, we hope to establish a relationship between the resource allocation

scheme and the accuracy of the neural network output. However, due to the current lack of understanding of the working mechanism of neural networks, it is challenging to establish such a relationship. To solve this problem, we propose a compromise solution. The output error of an effective UCFA is positively correlated with the error in the input cCSI and sCSI. Therefore, we only need to ensure that the channel estimation error or the radar sensing accuracy meets the minimum requirements, and we can approximate that the fCSI generated by the UCFA is the true channel.

4.1 Transmission Rate

Considering the uplink process, we employ the Minimum Mean Square Error (MMSE) algorithm for channel estimation. Consequently, all variables and matrices differ in uplink and downlink will be added the upper subscript as $\cdot^{(\mathrm{U})}$ or $\cdot^{(\mathrm{D})}$, and thus, the estimated channel matrix can be expressed as:

$$\hat{\mathbf{H}}_P^{(\mathrm{U})} = \mathbf{Y}_C^{(\mathrm{U})} \mathbf{X}_C^{(\mathrm{U}),\mathrm{H}} \left(\frac{P}{\sigma^2_{\mathbf{Z}_C^{(\mathrm{U})}} N_{\mathrm{U}}} \mathbf{I}_{N_{\mathrm{U}}} + \mathbf{X}_C^{(\mathrm{U}),\mathrm{H}} \mathbf{X}_C^{(\mathrm{U})} \right)^{-1}, \tag{6}$$

where $\mathbf{X}_C^{(\mathrm{U})} = \mathbf{F}_C^{(\mathrm{U})} \mathbf{S}_C^{(\mathrm{U})}$ denotes the transmitted signal after beamforming, $P = \mathrm{tr}(\mathbf{S}_C^{(\mathrm{U}),\mathrm{H}} \mathbf{S}_C^{(\mathrm{U})})$ is the transmission power and $\sigma^2_{\mathbf{Z}_C^{(\mathrm{U})}} = \mathrm{tr}(\mathbf{Z}_C^{(\mathrm{U}),\mathrm{H}} \mathbf{Z}_C^{(\mathrm{U})})$ is the power of environmental noise. Additionally, $\mathbf{I}_{N_{\mathrm{U}}}$ denotes the $N_{\mathrm{U}} \times N_{\mathrm{U}}$ identity matrix, while the other identity matrices are denoted similarly with the subscript as the order. Denoting the error of estimation channel as $\tilde{\mathbf{H}}_C^{(\mathrm{U})}$, the magnitude of the error matrix can be expressed as [12]:

$$\sigma^2_{\tilde{\mathbf{H}}_C^{(\mathrm{U})}} = \frac{\sigma^2_{\mathbf{Z}_C^{(\mathrm{U})}} \sigma^2_{\hat{\mathbf{H}}_C^{(\mathrm{U})}}}{\sigma^2_{\mathbf{Z}_C^{(\mathrm{U})}} + \frac{P_C^{(\mathrm{U})}}{N_{\mathrm{U}}} N_{\mathrm{P}}^{(\mathrm{U})} \sigma^2_{\hat{\mathbf{H}}_C^{(\mathrm{U})}}}, \tag{7}$$

where $\sigma^2_{\hat{\mathbf{H}}_C^{(\mathrm{U})}}$ is the variance of the estimated channel and can be obtained by the estimator in the BS. Besides, $N_{\mathrm{P}}^{(\mathrm{U})}$ means the number of pilot symbols in one transmission frame.

Although the fCSI in the uplink process is considered to be the true channel, errors exist in the channel obtained based on channel reciprocity in the downlink, so the downlink transmission rate should take into account the imperfect channel situation. In the design of beamforming, the system can only utilize the estimated channel rather than the true channel. Therefore, the error between the estimated and true channels should be regarded as a form of random interference, and it should be combined with environmental noise as the effective noise, $\mathbf{Z}_C^{(\mathrm{D})\prime}$. Thus, the received signal can be transformed as:

$$\mathbf{Y}_C^{(\mathrm{D})} = \hat{\mathbf{H}}_C^{(\mathrm{D})} \mathbf{F}_C^{(\mathrm{D})} \mathbf{S}_C^{(\mathrm{D})} + \underbrace{\tilde{\mathbf{H}}_C^{(\mathrm{D})} \mathbf{F}_C^{(\mathrm{D})} \mathbf{S}_C^{(\mathrm{D})} + \mathbf{Z}_C^{(\mathrm{D})}}_{\mathbf{Z}_C^{(\mathrm{D})\prime}}. \tag{8}$$

Therefore, the Shannon formula for a known channel can be modified to formula considering an imperfect channel. When only using channel estimation to obtain CSI, transmission rate can be expressed as:

$$\mathrm{R}_C^{(\mathrm{D})} = \log_2 \det(\mathbf{I} + \frac{P_C^{(\mathrm{D})}}{P_C^{(\mathrm{U})}\sigma^2_{\tilde{\mathbf{H}}_C^{(\mathrm{D})}} + \sigma^2_{\mathbf{Z}_C^{(\mathrm{D})}}}\mathbf{H}_C^{(\mathrm{D})}\mathbf{F}_C^{(\mathrm{D})}\mathbf{F}_C^{(\mathrm{D}),\mathrm{H}}\mathbf{H}_C^{(\mathrm{D}),\mathrm{H}}). \tag{9}$$

4.2 Sensing Accuracy

When there are M scatterers in the environment, there may be one LOS path, M single-hop NLOS paths, and several multi-hop NLOS paths between the user and the BS. In radar detection, we assume that the user and the BS have already shared the sensing results. The system can then select the common targets from both parties' detection results. For these common targets, the "user-target-BS" path exists. Otherwise, it implies that the path is partially blocked, or a false target has been generated due to a false alarm.

According the analysis, the error in radar-reconstructed channels primarily consists of three components: radar missed detections, the omission of multi-hop paths, and estimation errors of path gain coefficients. Correspondingly, the sensing system needs to ensure the probability of missed detection and the accuracy of the positioning parameter estimates. The probability of missed detection in the sensing system can be expressed as [13]:

$$P_M = (1 - Q_1\left(\sqrt{\frac{2P_S\mathrm{tr}(\mathbf{H}_S^{\mathrm{H}}\mathbf{H}_S)}{\sigma^2_{\mathbf{Z}_S}}}, \sqrt{2\gamma}\right))^{N_S}, \tag{10}$$

where $Q_1(a, b)$ denotes the first-order Marcum Q-function with a and b as variables. Besides, N_S is the number of pulse repetition, P_S is the power of the radar signal, $\mathbf{X}_{\mathrm{R}}$, and γ is the detection power threshold. If the detection threshold is relaxed, i.e., γ is small, the probability of detecting a target increases. Finally, in target localization, the Cramér-Rao Lower Bound (CRLB) for delay and angle is widely used to characterize localization accuracy [13]. They can be expressed as follows:

$$\mathrm{CRLB}_\tau = \frac{1}{8\pi^2 \mathrm{SNR} N_{\mathrm{U}} \beta^2 N_S}, \tag{11}$$

$$\mathrm{CRLB}_{\phi_{\mathrm{T}}} = \frac{3}{2\pi^2 \frac{d^2}{\lambda^2}\cos^2\phi_{\mathrm{T}} \mathrm{SNR} N_{\mathrm{U}}(N_{\mathrm{U}}^2 - 1)N_S}, \tag{12}$$

where $\beta^2 = \frac{\int_{-\infty}^{\infty} f^2|S(f)|^2 df - (\int_{-\infty}^{\infty} f|S(f)|^2 df)^2}{\int_{-\infty}^{\infty}|S(f)|^2 df}$ is the squared effective bandwidth, $S(f)$ is the FFT of transmitted baseband signal, $\mathbf{S}_S$, and the SNR is the received signal-to-noise ratio. Moreover, d and λ are the antenna spacing and signal wavelength, τ and ϕ_{T} are the estimated parameters, signal delay and angle of transmission, respectively.

4.3 Problem Formulation

We hope to optimize the resource allocation and beamforming to increase the transmission rate while satisfying the lowest sensing accuracy requirements. As analysis in Sect. 2.2, the ratio of transmission, uplink and downlink should keep constant, therefore, we focus on the power and uplink frame structure. In our system, the problem can be classified to uplink and downlink. The problem for uplink can be formulated as:

$$\mathcal{P}1: \max_{N_{\mathrm{P}}^{(\mathrm{U})}, N_{\mathrm{D}}^{(\mathrm{U})}, \mathbf{F}_C^{(\mathrm{U})}, P_S, P_C^{(\mathrm{U})}} N_{\mathrm{D}}^{(\mathrm{U})} \log_2 \det(\mathbf{I}_{N_B} + \frac{P_C^{(\mathrm{U})}}{\sigma^2_{\mathbf{Z}_C^{(\mathrm{U})}}} \mathbf{H}_C^{(\mathrm{U})} \mathbf{F}_C^{(\mathrm{U})} \mathbf{F}_C^{(\mathrm{U}),\mathrm{H}} \mathbf{H}_C^{(\mathrm{U}),\mathrm{H}}) \tag{13}$$

$$\begin{aligned}
s.t. \quad & \mathrm{CRLB}_{\tau,n} \le \mathrm{CRLB}_{\tau,\max}, \mathrm{CRLB}_{\phi_{\mathrm{T}},n} \le \mathrm{CRLB}_{\phi_{\mathrm{T}},\max}, \forall n \in \mathcal{M}, && \text{(13a)}\\
& P_{M,n} \le P_{M,\max}, \forall n \in \mathcal{M}, && \text{(13b)}\\
& N_{\mathrm{P}}^{(\mathrm{U})} + N_{\mathrm{D}}^{(\mathrm{U})} = N_C^{(\mathrm{U})}, N_{\mathrm{D}}^{(\mathrm{U})} \ge N_{\mathrm{D},\min}^{(\mathrm{U})} && \text{(13c)}\\
& N_S P_S + N_C^{(\mathrm{U})} P_C^{(\mathrm{U})} \le E_{\mathrm{total}}, && \text{(13d)}\\
& \mathrm{tr}(\mathbf{F}^{(\mathrm{U})} \mathbf{F}^{(\mathrm{U}),\mathrm{H}}) \le 1, && \text{(13e)}\\
& \sigma^2_{\hat{\mathbf{H}}_C^{(\mathrm{U})}} \le \sigma^2_{\hat{\mathbf{H}}_{C,\max}^{(\mathrm{U})}}, && \text{(13f)}
\end{aligned}$$

where $\mathrm{CRLB}_{\tau,\max}$ and $\mathrm{CRLB}_{\phi_{\mathrm{T}},\max}$ denote the maximum acceptable localization error of the system, and $P_{M,\max}$ is the requirement for probability of miss detection. All the sensing metrics in (13a) and (13b) with subscript n denote the metrics when detecting the n-th target. Besides, the constraint (13c) keeps that all frames contains the same number of symbols, while $N_{\mathrm{P}}^{(\mathrm{U})}, N_{\mathrm{D}}^{(\mathrm{U})}, N_C^{(\mathrm{U})}$ are numbers of pilot symbols, data symbols and total symbols. In (13d), N_S is the repetition number of sensing pulse signal, E_{total} denotes the total power which can be used in one cycle by the UE, and $P_S, P_C^{(\mathrm{U})}$ denote the power of sensing and communication respectively. Finally, (13e) limits the power of beamforming and (13f) limits the error of CE.

On the other hand, the downlink problem will be solve in the BS's equipment. The BS is equipped with dedicated devices for sensing and can utilize separate frequency bands. This is because the sensing tasks of the BS are typically more important than those of individual users. Therefore, in the downlink problem, the BS only needs to optimize the power allocation and beamforming for transmitting signals to multiple users. To maximize the overall transmission rate for all users, we design the optimization problem with the objective of maximizing the weighted sum rate of multiple users as follows:

$$\mathcal{P}2: \max_{\{\mathbf{F}_{C,i}^{(\mathrm{D})}\}, \{P_{C,i}^{(\mathrm{D})}\}} \sum_{i=1}^{K} \beta_i \mathrm{R}_{C,i}^{(\mathrm{D})} \tag{14}$$

$$\begin{aligned}
s.t. \quad & \mathrm{tr}(\mathbf{F}_{C,i}^{(\mathrm{D})} \mathbf{F}_{C,i}^{(\mathrm{D}),\mathrm{H}}) \le 1, \forall 1 \le i \le K, && \text{(14a)}\\
& \sum_{i=1}^{K} P_{C,i}^{(\mathrm{D})} = P_{\mathrm{sum}}, && \text{(14b)}
\end{aligned}$$

where $\mathbf{F}_{C,i}^{(\mathrm{D})}$ and $P_{C,i}^{(\mathrm{D})}$ are the beamforming and transmitting power for i-th user in downlink. Also, the i-th user's transmission rate in downlink, $P_{C,i}^{(\mathrm{D})}$, can refer to (9). Before optimization, the BS will distribute the weight, β_i, to i-th user and allocate the power according to these weights. Besides, the sum of allocated power is limited beyond P_{sum} in (14b).

5 The Resource Allocation Algorithm

In this section, we proprose two algorithms to solve the uplink and downlink problem, respectively.

5.1 Algorithm for Uplink

All the variables can be divided to two parts: beamforming, $\mathbf{F}_C^{(\mathrm{U})}$, and the others for resource allocation. If we fix the resource variables, the beamforming can be solved easily and fast by the traditional algorithm, such as water-filling method. Thus, we consider to simplify the resource variables at first. Considering the constraints (13a) and (13b), the CRLBs and the probabilities are negatively correlated with the sensing power. Therefore, given the requirements for sensing metrics, the lowest feasible power can be determined. For example, the lowest feasible power for CRLBs can expressed as:

$$P_{S,\min}^{(\mathrm{U})} = \min_{i\in\mathcal{M}} \{ \frac{\sigma_{\mathbf{Z}_S}^2 \alpha_i^{-1}}{8\pi^2 N_{\mathrm{U}} \beta^2 N_{\mathrm{PR}} \mathrm{CRLB}_{\tau,\max}}, \frac{3\sigma_{\mathbf{Z}_S}^2 \alpha_i^{-1}}{2\pi^2 \frac{d^2}{\lambda^2} \cos^2\phi_{\mathrm{T}} N_{\mathrm{U}} (N_{\mathrm{U}}^2 - 1) N_{\mathrm{PR}} \mathrm{CRLB}_{\phi_{\mathrm{T},\max}}} \}. \tag{15}$$

Similarly, for (13b), the minimum feasible sensing power can also be expressed as a function related to the number of P_S and certain indicators. According to constraints (13c) and (13d), both the highest feasible $P_C^{(\mathrm{U})}$ and $N_D^{(\mathrm{U})}$ can be represented as functions of P_S and $N_P^{(\mathrm{U})}$ as

$$P_{C,\max}^{(\mathrm{U})} = \frac{E_{\mathrm{total}} - N_S P_S}{N_C^{(\mathrm{U})}}, \tag{16}$$

$$N_{\mathrm{D}}^{(\mathrm{U})} = N_C^{(\mathrm{U})} - N_{\mathrm{P}}^{(\mathrm{U})}. \tag{17}$$

As far, all variables can be transformed to the expressions about $N_{\mathrm{P}}^{(\mathrm{U})}$ and P_S. This means we can search for the optimal power allocation scheme by traversing these two variables. For each search node, we calculate the beamforming using traditional methods and then select the optimal scheme among all nodes. Since the transmission rate is monotonically increasing with respect to power, we only need to perform a binary search on the power for each possible number of pilot symbols. The complexity of our algorithm is $\mathcal{O}(N_C^{(\mathrm{U})} \log(\frac{E_{\mathrm{total}}}{P_s}))$ where P_s is the minimum allocable energy unit.

5.2 Algorithm for Downlink

In the downlink, the $\mathcal{P}2$ is a convex problem. Therefore, we employ the block coordinate descent method, iteratively optimizing beamforming and power allocation. For optimizing beamforming, we utilize the water-filling algorithm, and for optimizing power, we apply the steepest gradient descent method. Assuming the gradient of the weighted transmission rate with respect to power as $\frac{\partial R_C^{(\mathrm{D})}}{\partial P_{C,i}^{(\mathrm{D})}}$, we can derive the iterative expression as follows:

$$P_{C,i,j+1}^{(\mathrm{D})} = P_{C,i,j}^{(\mathrm{D})} - \eta_j \frac{\partial R_C^{(\mathrm{D})}}{\partial P_{C,i}^{(\mathrm{D})}}, \tag{18}$$

where $P_{C,i,j}$ denotes the power allocated to i-th user in j-th iteration and η_j denotes the step in i-th iteration. For this alternative algorithm, the complexity depends on the more complex part of the two parts, which is the beamforming. Thus, the complexity of the algorithm depends on the complexity of the singular value decomposition in beamforming, that is, $\mathcal{O}(N_B^3)$.

6 Numerical Simulation

In this section, we evaluate the our UCFA and resource allocation algorithms. In our simulations, we assume the BS and users are equipped with 64 and 16 antennas working at a center frequency of $f_c = 30$ GHz. The TDD duration is 10 ms within 50 symbols. In a multi-user scenario, all users are placed in a long street where the buildings can scatter signals as shown in Fig. 1. Meanwhile, according to the urban channel model, the path loss is be given by:

$$PathLoss = c_1 + 10c_2 \log_{10}(d) + \xi(dB), \tag{19}$$

where $\xi \sim \mathcal{N}(0, \sigma_\xi^2)$ and c_1, c_2 differ in different cases. We adopt $c_1 = 61.4, c_2 = 2$ and $\sigma_\xi^2 = 5.8$ in LOS path, $c_1 = 72.0, c_2 = 2.92$ and $\sigma_\xi^2 = 8.7$ in NLOS paths. Besides, all obstacles are considered to reflect the omnidirectional signal.

For our UCFA, We create new datasets by randomly placing the users in the street. Subsequently, we construct the true channel and cCSI, sCSI by considering the random noise in CE and radar reconstruction.

As for basic setting for training, we follow the strategy of original Uformer and train the proposed neural network using the AdamW optimizer with the momentum terms of (0.9, 0.999) and the weight decay of 0.02 on a GeForce RTX 4050. Then, We use the cosine decay strategy to decrease the learning rate to 1e-6 with the initial learning rate 2e-4. We set the window size to 8×8 in all LeWin Transformer blocks. Then, We employ a cosine decay strategy to gradually decrease the learning rate from 2e-4 to 1e-6. Additionally, we set the window size to 8×8 in all LeWin Transformer blocks.

6.1 The Convergence and Accuracy Improvement of Uformer

As shown in the Fig. 4, we compare the Normalized Mean Square Error (NMSE) of our proposed algorithm and the traditional methods. In this case, we simulate 500 data and calculate cCSI, sCSI by traditional methods and fCSI by UCFA. The results of

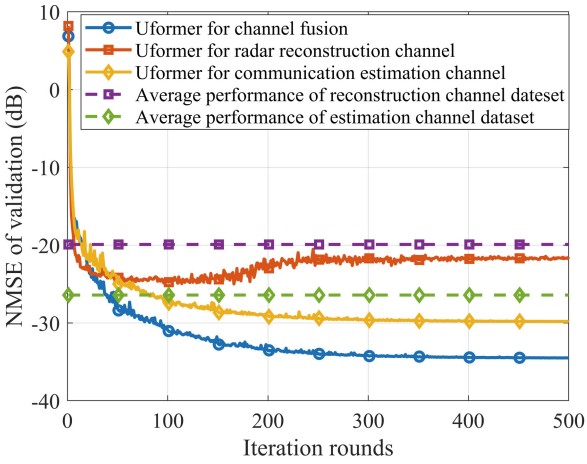

Fig. 4. The convergence and accuracy of the UCFA and baselines.

the traditional algorithms will be used as the input of the UCFA, and the method we proposed have achieved an error reduction of about 10dB compared with the traditional method. At the same time, in order to confirm that UCFA does combine the information of the dual input channels, we conducted an ablation experiment. The red and yellow lines are the network performances of only inputting sCSI and cCSI respectively, and the same channel will be copied as two inputs. Among them, the algorithm of inputting sCSI quickly turn overfitting due to lack of information, and as a result, our algorithm can perform best.

To show how our algorithm works, we visualize a set of data and present it in the Fig. 5. All channels are transformed using the FFT to display spatial characteristics. Specifically, (a), (b) and (c) represent the true channels, sCSI and cCSI, while (d), (e), and (f) are the output results of three different networks. Compared to (c), the sCSI in (b) does not have significant noise, but its estimation of small peaks is inaccurate because it struggles to capture precise amplitude information. After network processing, the channel with dual inputs in (d) achieves the best results, as it provides more accurate restoration of small peaks compared to (e), and demonstrate better overall denoising performance compared to (f).

6.2 The Comparison Between Our Optimization Scheme and Average Scheme

From the Fig. 6(a), it can be observed that as the sensing accuracy requirements increase, the transmission rate correspondingly decreases. This is because more energy is allocated to the sensing part. Secondly, when the system's available power increases, the rate shows a slight step increase during continuous improvement. This is due to the resource allocation algorithm assigning one symbol from the pilot to the data. As the total power increases, the number of symbols required for the pilot decreases. Each additional data symbol can increase the transmission rate by over 1%.

In the Fig. 6(b), we consider the power allocation to three users and the user with higher index are placed more close the BS. The transmission rates under average power

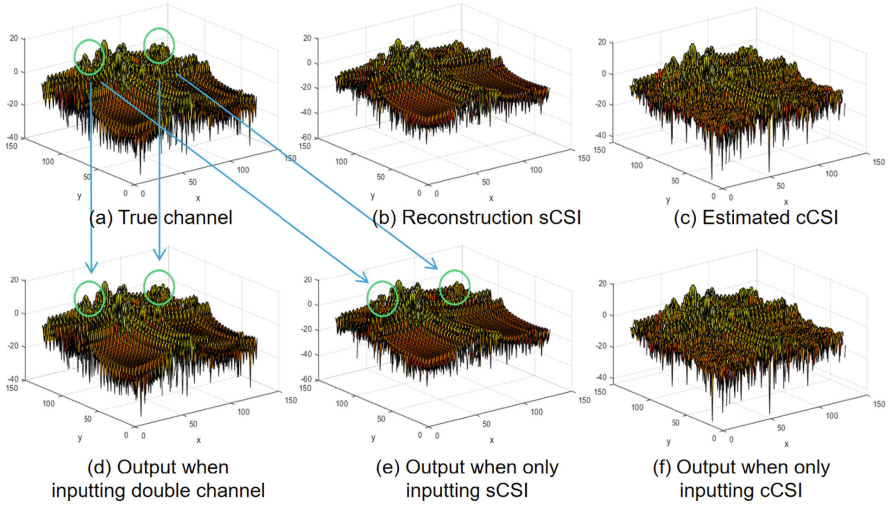

Fig. 5. An visualized example for three Uformer methods.

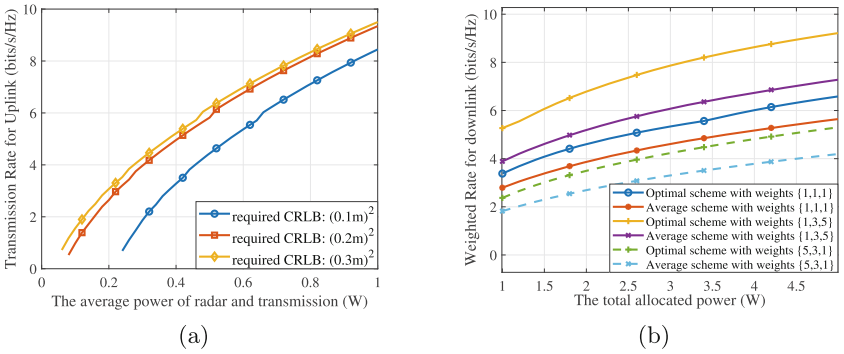

Fig. 6. The transmission rates of the resource allocation algorithms: (a) Uplink process under various sensing requirements; (b) Downlink process under three weights schemes.

allocation scheme and our optimal scheme are compared. It can be seen that our algorithm can always achieve higher rate.

7 Conclusion

This paper designs a novel channel fusion algorithm, which can combine the information in sensing data and communication CSI. Then, we design an ISAC system that allows the algorithm to participate in CE and enhance its accuracy. This system obtains sensing data and assists transmission by alternating between sensing and transmission. Based on this, we construct optimization problems for uplink and downlink to optimize resource allocation and maximize transmission rates while meeting sensing requirements. Additionally, we propose uplink and downlink resource optimization

algorithms. The performance of our designed UCFA and resource allocation algorithms is also validated through simulations.

Acknowledgments. This work is supported in part by the National Natural Science Foundation of China under Grant No. 62271438, 62062031, in part by Zhejiang Provincial Natural Science Foundation of China under Grant No. LGG22F010008, and in part by ROIS NII Open Collaborative Research under Grant 24S0601.

References

1. Zhu, X.: Integrated satellite-terrestrial networks toward 6G: architectures, applications, and challenges. IEEE Internet Things J. **9**(1), 437–461 (2021)
2. Giordani, M., Zorzi, M.: Satellite communication at millimeter waves: a key enabler of the 6G era. In: 2020 International Conference on Computing, Networking and Communications, pp. 383–388, (2020)
3. Heo, J., Sung, S.: MIMO satellite communication systems: a survey from the PHY layer perspective. IEEE Commun. Surv. Tutorials **25**(3), 1543–1570 (2023)
4. Ren, Z.: Fundamental CRB-rate tradeoff in multi-antenna ISAC systems with information multicasting and multi-target sensing. IEEE Trans. Wireless Commun. **23**(4), 3870–3885 (2024)
5. Hua, H., Xu, J.: Optimal transmit beamforming for integrated sensing and communication. IEEE Trans. Veh. Technol. **72**(8), 10588–10603 (2023)
6. Mundlamuri, R.: Sensing aided channel estimation in wideband millimeter-wave MIMO systems. In: 2023 IEEE International Conference on Communications Workshops, pp. 1404–1409. (2023)
7. Fan, L.: Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond. IEEE J. Sel. Areas Commun. **40**(6), 1728–1767 (2022)
8. Jiang, S., Alkhateeb, A.: Sensing aided OTFS massive MIMO systems: compressive channel estimation. In: 2023 IEEE International Conference on Communications Workshops, pp. 794–799, (2023)
9. Chen, X., Feng, Z.: Sensing-aided uplink channel estimation for joint communication and sensing. IEEE Wireless Commun. Lett. **12**(3), 441–445 (2023)
10. Jiajia, G.: Deep learning for joint channel estimation and feedback in massive MIMO systems. Digital Commun. Netw. **10**(1), 83–93 (2024)
11. Wang, U.: Uformer: a general u-shaped transformer for image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17683–17693 (2022)
12. Hassibi, B.: Bertr: How much training is needed in multiple-antenna wireless links? IEEE Trans. Inf. Theory **49**(4), 951–963 (2003)
13. Liu, A.: A survey on fundamental limits of integrated sensing and communication. IEEE Commun. Surv. Tutorials **24**(2), 994–1034 (2022)