

Joint Localization and Communication Enhancement in Uplink Integrated Sensing and Communications System with Clock Asynchronism

Xu Chen *, *Member, IEEE*, XinXin He *, *Member, IEEE*, Zhiyong Feng, *Senior Member, IEEE*, Zhiqing Wei, *Member, IEEE*, Qixun Zhang, *Member, IEEE*, Xin Yuan, *Member, IEEE*, and Ping Zhang, *Fellow, IEEE*

Abstract—In this paper, we propose a joint single-base localization and communication enhancement scheme for the uplink (UL) integrated sensing and communications (ISAC) system with asynchronism, which can achieve accurate single-base localization of user equipment (UE) and significantly improve the communication reliability despite the existence of timing offset (TO) due to the clock asynchronism between UE and base station (BS). Our proposed scheme integrates the CSI enhancement into the multiple signal classification (MUSIC)-based AoA estimation and thus imposes no extra complexity on the ISAC system. We further exploit a MUSIC-based range estimation method and prove that it can suppress the time-varying TO-related phase terms. Exploiting the AoA and range estimation of UE, we can estimate the location of UE. Finally, we propose a joint CSI and data signals-based localization scheme that can coherently exploit the data and the CSI signals to improve the AoA and range estimation, which further enhances the single-base localization of UE. The extensive simulation results show that the enhanced CSI can achieve equivalent bit error rate performance to the minimum mean square error (MMSE) CSI estimator. The proposed joint CSI and data signals-based localization scheme can achieve decimeter-level localization accuracy despite the existing clock asynchronism and improve the localization mean square error (MSE) by about 8 dB compared with the maximum likelihood (ML)-based benchmark method.

Index Terms—Joint communications and sensing (JCAS), integrated sensing and communications (ISAC), uplink localization, CSI enhancement, 6G.

I. INTRODUCTION

A. Backgrounds and Motivations

With the rapid development of the sixth-generation (6G) intelligent machine-type applications, such as intelligent vehicular networks, smart factories, and smart cities, both communication and sensing are indispensable for the flexible

This work is supported by the National Key Research and Development Program of China under Grants {2020YFA0711300, 2020YFA0711302, and 2020YFA0711303}.

*These authors contributed equally to this work and should be regarded as co-first authors.

Xu Chen, XinXin He, Z. Feng, Zhiqing Wei, and Qixun Zhang are with School of Information and communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, P. R. China (Email:{chenxu96330, hxx_9000, fengzy, weizhiqing, zhangqixun}@bupt.edu.cn).

X. Yuan is with the University of Technology Sydney, Ultimo, NSW 2007, Australia (email: xin.yuan@ieee.org).

Ping Zhang is with State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, P. R. China (Email: pzhang@bupt.edu.cn).

Corresponding author: Zhiyong Feng

operation of autonomous machines [1]. Integrated sensing and communications (ISAC), also known as joint communication and sensing (JCAS), is a promising technique to solve the spectrum congestion problem due to the proliferation of wireless communication and sensing devices [2], [3]. ISAC can share the same transceivers and spectrum to achieve both sensing and communication functions using the same transmitted signals [1], [4], [5].

ISAC achieved by using uplink (UL) signals, i.e., UL ISAC, has become a promising approach to deploying ISAC techniques in the cellular communication systems [6], since it does not require full-duplex (FD) operations as in the downlink (DL) ISAC [7]. One example of UL ISAC in communication systems is to conduct sensing at the base station (BS) using the UL communication signals from user equipment (UE). Consequently, UL ISAC can be implemented with minimal adjustments to the network infrastructure. Nonetheless, the disparity in clock synchronization between the BS and UE introduces a fluctuating timing offset (TO), posing a substantial obstacle to achieving precise range and location estimations in UL ISAC. Therefore, mitigating the sensing ambiguity from clock asynchronism is one of the biggest challenges in UL ISAC.

In this paper, we focus on the single-base localization of UE in UL ISAC, which does not require solving the complex clock synchronization issues among multiple BSs as conducted in the multi-base cooperative localization [8]. Moreover, as illustrated in the early works [9], the angle-of-arrival (AoA) estimation in UL ISAC is not prominently affected by the clock asynchronism. Therefore, using the AoA and range estimation of a single BS to localize the UE can avoid the partial influence of sensing ambiguity. Furthermore, we consider enhancing the channel state information (CSI) estimation in the process of AoA estimation in UL ISAC to improve communication reliability.

B. Related Works

In this subsection, we review the existing works related to the UL ISAC localization and CSI estimation.

For a long time, cooperation among multiple BSs and iterative location estimation using multiple packets has been prevalent to deal with the range and location offset due to TO. In [10], the authors proposed an expectation-maximization-based cooperative localization method. This method requires

numerous receivers and more than 20 iterations for expectation maximization, which results in a high implementation complexity. In [8], the authors proposed a cooperative beamforming and power allocation scheme with multiple BSs to maximize the maximum rate and positioning error bounds. However, this paper does not present the implementable signal processing method for achieving the localization performance bounds and ignores the influence of TO.

Recently, new techniques based on the feature of multiple antennas have been proposed to address the clock asynchronism problem. In [11], the authors proposed to use the cross-antenna cross-correlation (CACC) method to track humans passively with a single WiFi link by exploiting the cross-correlation among all the pairs of antennas. In [12], the authors proposed a UL ISAC method for perceptive mobile networks, allowing a static UE and BS to form a bi-static system to sense the environment by utilizing the CACC method. However, this method only works under the assumption that the accurate location of UE is known in advance and has to solve the challenging image target problems, which means it is unsuitable for localizing UEs. In [13]–[15], the authors proposed to use the cross-antenna signal ratio (CASR) method to estimate the Doppler frequency exploiting the CSI ratio between each two antennas. However, this method works only when the scatterers are static except for the target of interest to maintain the linearity of the sensing parameter estimation problem based on the CSI ratio.

On the other hand, the CSI estimators and enhancers are also promising research areas in ISAC. The least-square (LS) method is widely used to estimate CSI for communication demodulation and sensing parameter estimation [16], [17] due to its low complexity. In [18], the authors proposed to use the sensing parameter estimates to enhance the CSI estimation accuracy for improving communication reliability. In [19], the authors proposed a Kalman filter (KF)-based communication CSI enhancer by exploiting the AoA estimates to suppress the random noise terms in LS CSI to improve the communication performance. However, these CSI enhancers require adding filtering procedures into the original ISAC signal processing, which imposes extra complexity.

C. Contributions

In this paper, we propose a novel joint single-base localization and communication enhancement scheme for UL ISAC. We first propose a joint AoA estimation and CSI enhancement method, which integrates the CSI enhancement into the multiple signal classification (MUSIC)-based AoA estimation procedures. Unlike the existing CSI enhancers, this method is completed in the AoA estimation and imposes no extra complexity on the ISAC system. Subsequently, we further reveal that the MUSIC-based range estimation method can suppress the TO-related noise-like phase terms in CSI estimation, which can reduce the range ambiguity. Using the AoA and range estimation, we can estimate the location of UE with spherical coordinates. Finally, we propose a novel joint CSI and data signals-based localization scheme that uses the combined demodulated data and CSI signals to improve the sensing performance.

The main contributions of this paper are summarized as follows.

1. We propose a joint AoA estimation and CSI enhancement method that uses the eigenvalue decomposition procedure in the MUSIC-based AoA estimation to suppress the noise terms in the initial CSI estimation obtained by LS estimator. This method requires no additional filters to enhance the CSI estimation and imposes no extra complexity to the ISAC system. The simulation results show that the proposed joint AoA estimation and CSI enhancement method can achieve the same bit error rate (BER) performance as the minimum mean square error (MMSE) estimator.
2. We reveal that the MUSIC-based range estimation method can suppress the TO-related time-varying noise terms by exploiting the time-domain averaging effects on multiple packets in the eigenvalue decomposition procedures. Moreover, the range estimation of UE is not affected by the noise and offset terms in frequency and antenna domains since the frequency and antenna-domain noise is coherently absorbed in the eigenvalue decomposition procedures.
3. We propose a novel joint CSI and data signals-based localization scheme that can flexibly combine the CSI and demodulated data signals to improve the AoA and range estimation, which naturally improves the localization performance. We prove that the MUSIC-based AoA and range estimation can coherently utilize the data signals to improve the AoA and range estimation.

We provide extensive simulation results, validating the proposed joint single-base localization and communication enhancement scheme. The localization mean square error (MSE) of the proposed scheme is shown to be about 8 dB lower than the maximum-likelihood (ML)-based benchmark scheme.

D. Outline of This Paper

The remaining parts of this paper are organized as follows. In section II, we describe the system model of the UL JCAS scheme. Section III proposes the joint single-base localization and CSI enhancement scheme. Section IV proposes the joint CSI and data signals-based localization scheme and analyzes its complexity. In section V, the simulation results are presented to validate the proposed scheme. Section VI concludes this paper.

Notations: Bold uppercase letters denote matrices (e.g., \mathbf{M}); bold lowercase letters denote column vectors (e.g., \mathbf{v}); scalars are denoted by normal font (e.g., γ); the entries of vectors or matrices are referred to with square brackets, for instance, the q th entry of vector \mathbf{v} is $[\mathbf{v}]_q$, and the entry of the matrix \mathbf{M} at the m th row and q th column is $[\mathbf{M}]_{n,m}$; $\mathbf{U}_s = [\mathbf{U}]_{:,N_1:N_2}$ means the matrices sliced from the N_1 th to the N_2 th columns of \mathbf{U} ; $(\cdot)^H$, $(\cdot)^*$ and $(\cdot)^T$ denote Hermitian transpose, complex conjugate and transpose, respectively; $\text{vec}(\cdot)$ is the operation to vectorize a matrix; $\mathbf{M}_1 \in \mathbb{C}^{M \times N}$ and $\mathbf{M}_2 \in \mathbb{R}^{M \times N}$ are $M \times N$ complex-value and real-value matrices, respectively; $\|\mathbf{v}\|_l$ represents the l -norm of \mathbf{v} , and ℓ_2 -norm is considered in

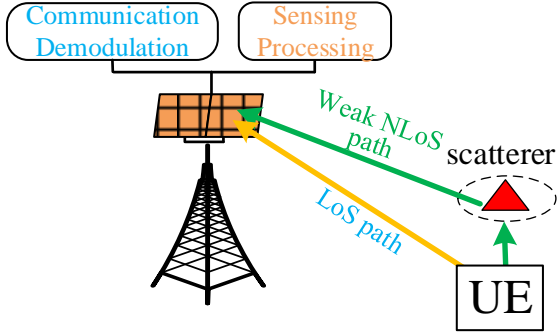


Fig. 1: The UL ISAC scenario.

this paper; for two given matrices \mathbf{S}_1 and \mathbf{S}_2 , $[v_{p,q}]_{(p,q) \in \mathbf{S}_1 \times \mathbf{S}_2}$ denotes the vector stacked by values $v_{p,q}$ satisfying $p \in \mathbf{S}_1$ and $q \in \mathbf{S}_2$; and $v \sim \mathcal{CN}(m, \sigma^2)$ means v follows a circular symmetric complex Gaussian (CSCG) distribution with mean value m and variance σ^2 .

II. SYSTEM MODEL

This section presents the UL ISAC system setup, channel models, and received signal models to provide fundamentals for UL ISAC signal processing.

A. UL ISAC System Setup

We consider a UL ISAC system where the BS and UE use uniform plane arrays (UPAs), as shown in Fig. 1. In the UL preamble period, UE transmits the pilot signals to BS, and BS uses the received training sequences for CSI estimation. In the UL data period, UE transmits the data signals to BS, and BS demodulates the UL data signals of UE using the estimated CSI.

In this paper, the CSI estimation and data symbols are both employed to enhance the estimation of the AoA and range to improve the localization performance. We assume that the pilot sequences for channel estimation are transmitted at an equal interval for simplicity of presentation. Each training sequence can be part of a typical packet in, e.g., WiFi systems, or transmitted in timeslot regularly, in, e.g., 5G mobile networks. There is one channel estimate during each packet. We also assume that within a coherent processing interval, M_s CSI estimates are obtained at an interval of T_s^p . The key idea in this paper can be extended to more general cases of non-uniform intervals.

Orthogonal frequency division multiplexing (OFDM) signal is utilized in the ISAC system. The key parameters for the OFDM signal are denoted as follows. P_t^U is the transmit power, N_c is the number of subcarriers occupied by UE in each OFDM symbol; f_c is the carrier frequency, T_s is the time duration of each symbol, and Δf is the subcarrier interval. Since there is one CSI estimate in each packet, $T_s^p = P_s T_s$, where P_s is the number of OFDM symbols in each packet.

B. UPA Model

The size of UPA is $P \times Q$, and we use d_a to denote the uniform interval between the neighboring antenna elements. The AoA for receiving or the angle-of-departure (AoD) for

transmitting the k th far-field signal is $\mathbf{p}_k = (\varphi_k, \theta_k)^T$, where φ_k and θ_k are the azimuth and elevation angles, respectively. The phase difference between the (p,q) th and the reference antenna elements is [20]

$$a_{p,q}(\mathbf{p}_k) = \exp[-j \frac{2\pi}{\lambda} d_a (p \cos \varphi_k \sin \theta_k + q \sin \varphi_k \sin \theta_k)], \quad (1)$$

where $\lambda = c/f_c$ is the wavelength of the carrier, f_c is the carrier frequency, and c is the velocity of light in vacuum. The steering vector for the array is given by

$$\mathbf{a}(\mathbf{p}_k) = [a_{p,q}(\mathbf{p}_k)]_{|p=0,1,\dots,P-1;q=0,1,\dots,Q-1}, \quad (2)$$

where $\mathbf{a}(\mathbf{p}_k) \in \mathbb{C}^{PQ \times 1}$. The sizes of the antenna arrays of the user and BS are $P_r \times Q_r$ and $P_t \times Q_t$, respectively.

C. UL ISAC Channel Model

In this paper, we assume there exists a strong LoS path and several weak reflective multipaths in the mmWave channel between UE and BS. The range and AoA of UE contained in the channel are also the sensing parameters to be estimated for localization. The UL ISAC channel response at the n th subcarrier of the m th packet can be expressed as [6]

$$\mathbf{H}_{n,m} = \sum_{k=0}^{K-1} \left[b_{C,k} e^{j2\pi f_{d,k} m T_s^p} e^{-j2\pi n \Delta f \tau_k} \times \mathbf{a}(\mathbf{p}_{R,k}^U) \mathbf{a}^T(\mathbf{p}_{T,k}^U) \right], \quad (3)$$

where $\mathbf{H}_{n,m} \in \mathbb{C}^{P_t Q_t \times P_r Q_r}$, $k=0$ is for the channel response of the LoS path, and $k \in \{1, \dots, K-1\}$ is for the k th non-line-of-sight (NLoS) path; $\mathbf{p}_{R,k}^U$ and $\mathbf{p}_{T,k}^U$ are the corresponding AoA and AoD, respectively; $\mathbf{a}(\mathbf{p}_{R,k}^U) \in \mathbb{C}^{P_t Q_t \times 1}$ and $\mathbf{a}(\mathbf{p}_{T,k}^U) \in \mathbb{C}^{P_r Q_r \times 1}$ are the steering vectors for UL reception and transmission, respectively; $f_{d,0} = \frac{v_{0,1}}{\lambda}$ and $\tau_{c,0} = \frac{r_{0,1}}{c}$ are the Doppler shift and time delay between UE and BS of the LoS path, respectively, with $v_{0,1}$ and $r_{0,1}$ being the corresponding radial relative velocity and the distance, respectively; $f_{d,k} = f_{d,k,1} + f_{d,k,2}$ and $\tau_k = \tau_{k,1} + \tau_{k,2}$ are the aggregate Doppler shift and time delay of the k th NLoS path, respectively; $f_{d,k,1} = \frac{v_{k,1}}{\lambda}$ and $f_{d,k,2} = \frac{v_{k,2}}{\lambda}$ are the Doppler frequencies between UE and the k th scatterer, and between the k th scatterer and the BS, respectively, with $v_{k,1}$ and $v_{k,2}$ being the corresponding radial velocities; $\tau_{k,1} = \frac{r_{k,1}}{c}$ and $\tau_{k,2} = \frac{r_{k,2}}{c}$ are the time delays between UE and the k th scatterer, and between BS and the k th scatterer, respectively, with $r_{k,1}$ and $r_{k,2}$ being the corresponding distances. Moreover, $b_{C,0} = \sqrt{\frac{\lambda^2}{(4\pi r_{0,1})^2}}$ and $b_{C,k} = \sqrt{\frac{\lambda^2}{(4\pi)^3 r_{k,1}^2 r_{k,2}^2}} \beta_{C,k}$ are the attenuations of the LoS and NLoS paths, respectively, where $\beta_{C,k}$ is the reflecting factor of the k th scatterer, following $\mathcal{CN}(0, \sigma_{C\beta,k}^2)$ [21].

D. CSI estimation

The CSI can be initially estimated based on, e.g., the LS method using the training sequences in the preamble [22]. We assume that a transmit beamforming (BF) with the BF vector \mathbf{w}_T is used. Omitting the details of training sequences and

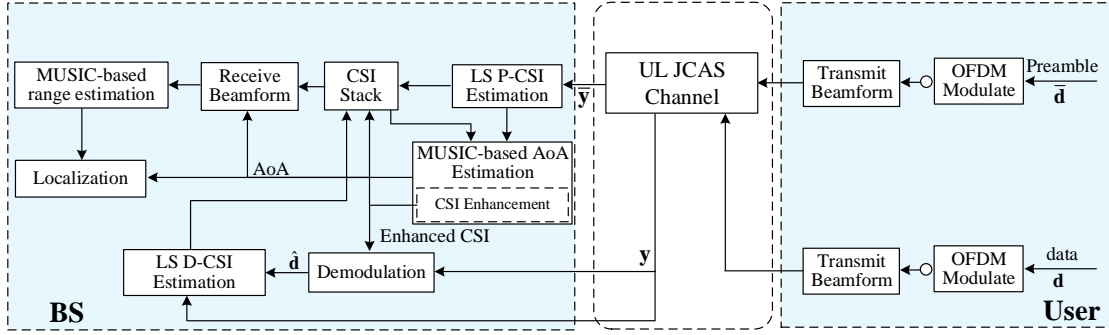


Fig. 2: The diagram for joint single-based localization and communication enhancement.

estimation method, we present the UL CSI estimated by the LS method at the n th subcarrier of the m th packet as

$$\begin{aligned} \hat{\mathbf{h}}_{n,m} &= \mathbf{h}_{n,m} + \mathbf{n}_{n,m} \in \mathbb{C}^{P_t Q_t \times 1} \\ &= \sqrt{P_t^U} \mathbf{H}_{n,m} \mathbf{w}_T \zeta_{f,m} \zeta_{\tau,m} + \mathbf{n}_{n,m} \\ &= \sqrt{P_t^U} \sum_{k=0}^{K-1} \begin{bmatrix} b_{C,k} e^{j2\pi m T_s^p [f_{d,k} + \delta_f(m)]} \\ \times e^{-j2\pi n \Delta f [\tau_k + \delta_\tau(m)]} \\ \times \chi_{T,k} \mathbf{a}(\mathbf{p}_{R,k}^U) \end{bmatrix} + \mathbf{n}_{n,m}, \end{aligned} \quad (4)$$

where $\mathbf{h}_{n,m} = \sqrt{P_t^U} \mathbf{H}_{n,m} \mathbf{w}_T \zeta_{f,m} \zeta_{\tau,m}$ is the equivalent channel response, $\zeta_{f,m} = e^{j2\pi m T_s^p \delta_f(m)}$ and $\zeta_{\tau,m} = e^{-j2\pi n \Delta f \delta_\tau(m)}$ are the phase shifts due to carrier frequency offset (CFO) and TO, denoted by $\delta_f(m)$ and $\delta_\tau(m)$, respectively; $\delta_f(m)$ and $\delta_\tau(m)$ are random time-varying parameters, following Gaussian distribution with zero mean and variances σ_f^2 and σ_τ^2 , respectively; T_s^p is the time interval between two CSI estimates; $\mathbf{n}_{n,m} \in \mathbb{C}^{P_t Q_t \times 1}$ is the combined noise that contains complex Gaussian noise, and each element of $\mathbf{n}_{n,m}$ follows $\mathcal{CN}(0, \sigma_N^2)$; \mathbf{w}_T is the transmit BF vector with $\|\mathbf{w}_T\|_2 = 1$, and $\chi_{T,k} = \mathbf{a}^T(\mathbf{p}_{T,k}^U) \mathbf{w}_T$ is the transmit BF gain.

E. Received Data Signals

The i th received OFDM symbol at the n th subcarrier of the m th packet can be expressed as

$$\mathbf{y}_{n,m}^i = \mathbf{h}_{n,m} d_{n,m}^i + \mathbf{n}_{n,m}^i \in \mathbb{C}^{P_t Q_t \times 1}, \quad (5)$$

where $d_{n,m}^i$ is the i th transmitted OFDM data symbol at the n th subcarrier of the m th packet, $\mathbf{n}_{n,m}^i$ is the Gaussian noise vector with each element following $\mathcal{CN}(0, \sigma_N^2)$, and $\mathbf{h}_{n,m}$ is the equivalent channel response in (4).

III. JOINT SINGLE-BASE LOCALIZATION AND CSI ENHANCEMENT IN PREAMBLE PERIOD

We present the UL single-base joint localization and communication enhancement scheme as shown in Fig. 2. This section introduces the signal processing procedures using the preamble signals. BS conducts initial preamble-based CSI (P-CSI) estimation using, e.g., the LS channel estimation method. Based on the singular value decomposition (SVD) of the estimated P-CSIs, we can estimate AoAs without being impacted by TO and CFO using a refined 2D MUSIC-based AoA estimation method and improve the accuracy of CSI estimation in the process of AoA estimation. Next, using the AoA estimates, we construct a spatial filter to separate CSIs with different AoAs. We then introduce the MUSIC-based

Algorithm 1: 2D two-step Newton descent minimum searching method [18]

Input: The range of φ : Φ_φ ; the range of θ : Φ_θ ; the number of grid points: N_i ; the maximum iteration round ind_{max} ; the MUSIC spectrum function: $f(\mathbf{p})$.

Output: Estimation results: $\Theta = \{\hat{\mathbf{p}}_k\}_{k \in \{0, \dots, N_s - 1\}}$.

Initialize:

1) Φ_φ and Φ_θ are both divided evenly into $N_i - 1$ pieces with N_i grid points to generate grid $\hat{\Phi}_\varphi$ and $\hat{\Phi}_\theta$.

2) Set a null space Θ .

Process:

Step 1: **foreach** $\mathbf{p}_{i,j} \in \hat{\Phi}_\varphi \times \hat{\Phi}_\theta$ **do**

 Calculate the spatial spectrum as \mathbf{S} , where
 $[\mathbf{S}]_{i,j} = [f(\mathbf{p}_{i,j})]^{-1}$.

end

Step 2: Search the maximal values of \mathbf{S} to form the set $\bar{\Theta}_d$.

Step 3: Derive the Hessian matrix and the gradient vector of $f(\mathbf{p})$ as $\mathbf{H}_p(\mathbf{p})$ and $\nabla_p f(\mathbf{p})$, respectively.

Step 4: **foreach** $\mathbf{p}_{i,j} \in \bar{\Theta}_d$ **do**

$k=1$;

$\mathbf{p}^{(0)} = \mathbf{p}_{i,j}$;

$\mathbf{p}^{(k)} = \mathbf{p}^{(k-1)} - [\mathbf{H}_p(\mathbf{p}^{(k-1)})]^{-1} \nabla_p f(\mathbf{p}^{(k-1)})$;

while $\|\mathbf{p}^{(k)} - \mathbf{p}^{(k-1)}\| > \varepsilon$ **and** $k \leq ind_{max}$ **do**

$\mathbf{p}^{(k)} = \mathbf{p}^{(k-1)} - [\mathbf{H}_p(\mathbf{p}^{(k-1)})]^{-1} \nabla_p f(\mathbf{p}^{(k-1)})$;

end

$\mathbf{p}^{(k)}$ is put into output set Θ ;

end

range estimation method to estimate the ranges exploiting the separated CSIs corresponding to the estimated AoAs. We prove that the MUSIC-based range estimation method can suppress the noise-like TO terms in the CSI to enhance the accuracy of range estimation. Finally, we can localize the UE with the spherical coordinates determined by the estimated AoA and range.

A. Joint AoA Estimation and CSI Enhancement

We choose the MUSIC-based method to estimate AoAs as it can work with non-equally spaced measurements in time, frequency, and spatial domains, offering great flexibility

to resource allocation. Such non-regular measurements are common in communication systems.

1) *AoA Estimation*: Stacking all $M_s \times N_c$ CSI estimates (from N_c subcarriers and M_s packets), we obtain the CSI matrix as

$$\hat{\mathbf{H}}_C = \mathbf{H} + \hat{\mathbf{N}} \in \mathbb{C}^{P_t Q_t \times N_c M_s}, \quad (6)$$

where \mathbf{H} is the actual value corresponding to $\hat{\mathbf{H}}_C$, and the $[(m-1)N_c + n]$ th columns of $\hat{\mathbf{H}}_C$, \mathbf{H} , and $\hat{\mathbf{N}}$ are $\hat{\mathbf{h}}_{n,m}$, $\mathbf{h}_{n,m}$, and $\mathbf{n}_{n,m}$, respectively.

Construct the correlation matrix of $\hat{\mathbf{H}}_C$ as

$$\mathbf{R}_x = [\hat{\mathbf{H}}_C (\hat{\mathbf{H}}_C)^H] / (M_s N_c). \quad (7)$$

Note that $\hat{\mathbf{H}}_C (\hat{\mathbf{H}}_C)^H = \sum_{n,m}^{N_c, M_s} \hat{\mathbf{h}}_{n,m} (\hat{\mathbf{h}}_{n,m})^H$. According to the proof in [9], the AoA estimation is not prominently affected by the CFO and TO, which will also be shown in Section V. Next, we use a refined 2D MUSIC method [18] to estimate the 2D AoA based on the SVD of $\hat{\mathbf{H}}_C$. The SVD of $\hat{\mathbf{H}}_C$ can be expressed as

$$\hat{\mathbf{H}}_C = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H = [\mathbf{U}_s, \mathbf{U}_0] \begin{bmatrix} \mathbf{\Sigma}_s & 0 \\ 0 & \mathbf{\Sigma}_0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_s^H \\ \mathbf{V}_0^H \end{bmatrix}, \quad (8)$$

where \mathbf{U}_s and \mathbf{U}_0 are the block matrices in the left singular matrix that are correlated to the signal and noise subspace, respectively; \mathbf{V}_s and \mathbf{V}_0 are the block matrices in the right singular matrix that are correlated to the signal and noise subspace, respectively; and $\mathbf{\Sigma}_s$ and $\mathbf{\Sigma}_0$ are the diagonal matrices that are composed of the signal and noise singular values, respectively. The number of the identified AoAs is the rank of $\mathbf{\Sigma}_s$ and can be ascertained by estimating the rank of the signal space from the differential vector derived from the singular values (or eigenvalues) of the signal matrix, which is presented in the appendix in [18].

We use N_A to denote the number of estimated AoAs. Then, the noise subspace of $\hat{\mathbf{H}}_C$ is $\mathbf{U}_0 = [\mathbf{U}]_{:,N_A+1:P_t Q_t}$, and we can formulate the angle spectrum function as

$$f_a(\mathbf{p}) = \mathbf{a}^H(\mathbf{p}) \mathbf{U}_0 (\mathbf{U}_0)^H \mathbf{a}(\mathbf{p}), \quad (9)$$

where $\mathbf{a}(\mathbf{p})$ is given in (2). The minimum points of $f_a(\mathbf{p})$ correspond to the AoAs to be estimated. We exploit a 2D two-step Newton descent method [18] to obtain the minimum points of $f_a(\mathbf{p})$, which is shown in **Algorithm 1**. To identify the minimum of $f_a(\mathbf{p})$, we substitute $f(\mathbf{p})$, $\mathbf{H}_p(\mathbf{p})$, and $\nabla_p f(\mathbf{p})$ in **Algorithm 1** with (9), Hessian matrix and the gradient vector of $f_a(\mathbf{p})$, respectively. Since the gain of LoS path is significantly larger than those of the weak NLoS paths, the estimated angle with the largest $f_a^{-1}(\mathbf{p})$ should be the AoA of UE, which is denoted by $\hat{\mathbf{p}}_{R,0}^U = (\hat{\varphi}_0, \hat{\theta}_0)$.

Based on the eigenvalue vector, denoted by $\mathbf{v}_s = \text{vec}(\mathbf{\Sigma}_s)$, obtained in the MUSIC process, we can also estimate the variance of $\hat{\mathbf{h}}_{n,m}$ as $\hat{\sigma}_N^2$. According to [23], $\mathbf{v}_s \in \mathbb{C}^{P_t Q_t \times 1}$ can be expressed as

$$[\mathbf{v}_s]_i = \begin{cases} P_i + \sigma_N^2, & \text{for } i \leq N_A, \\ \sigma_N^2, & \text{for } i > N_A, \end{cases} \quad (10)$$

where P_i is the power of the i th incident signal. Therefore, we obtain the estimation of the noise power as

$$\hat{\sigma}_N^2 = \sum_{i=N_A+1}^{P_t Q_t} [\mathbf{v}_s]_i / (P_t Q_t - N_A). \quad (11)$$

2) *CSI Enhancement*: Since the effective rank of $\hat{\mathbf{H}}_C$ is estimated as N_A , we obtain $\mathbf{\Sigma}_s = [\mathbf{\Sigma}]_{1:N_A,1:N_A}$, $\mathbf{U}_s = [\mathbf{U}]_{1:N_A,:}$, and $\mathbf{V}_s = [\mathbf{V}]_{1:N_A,:}$. The enhanced CSI estimation that eliminates the noise subspace can be expressed as

$$\bar{\mathbf{H}} = \mathbf{U}_s \mathbf{\Sigma}_s \mathbf{V}_s^H = \mathbf{H} + \bar{\mathbf{N}}, \quad (12)$$

where $\bar{\mathbf{N}}$ is the colored noise blended into the signal subspace. Here, $\bar{\mathbf{N}}$ has the following features. First, $\bar{\mathbf{N}}$ is irrelevant to \mathbf{H} , i.e., $E\{\mathbf{H}\bar{\mathbf{N}}^H\} = \mathbf{0}$. Second, the eigenvalues of the autocorrelation of $\bar{\mathbf{N}}$ are highly centralized, and there is $E\{\bar{\mathbf{N}}\bar{\mathbf{N}}^H\} \approx \mathbf{U}_s \sigma_N^2 \mathbf{U}_s^H$ [24]. Based on the above two features, we can obtain the autocorrelation of $\bar{\mathbf{H}}$ and \mathbf{H} , respectively, as

$$\mathbf{R}_{\bar{\mathbf{H}}} = E\{\bar{\mathbf{H}}\bar{\mathbf{H}}^H\} = \mathbf{U}_s \mathbf{\Sigma}_s^2 \mathbf{U}_s^H, \quad (13)$$

and

$$\begin{aligned} \mathbf{R}_{\mathbf{H}} &= E\{\mathbf{H}\mathbf{H}^H\} \\ &= \mathbf{R}_{\bar{\mathbf{H}}} - E\{\bar{\mathbf{N}}\bar{\mathbf{N}}^H\} \\ &= \mathbf{U}_s \mathbf{\Sigma}_s^2 \mathbf{U}_s^H - \mathbf{U}_s \sigma_N^2 \mathbf{U}_s^H \\ &= \mathbf{U}_s (\mathbf{\Sigma}_s^2 - \sigma_N^2 \mathbf{I}) \mathbf{U}_s^H \end{aligned} \quad (14)$$

Next, we form the following optimization problem to find the optimal combination of $\bar{\mathbf{H}}$, denoted by $\bar{\mathbf{H}}_u = \mathbf{B}\bar{\mathbf{H}}$, to suppress the colored noise in $\bar{\mathbf{H}}$.

$$\begin{aligned} \mathbf{B} &= \arg \min_{\mathbf{B}} E \|\mathbf{B}\bar{\mathbf{H}} - \mathbf{H}\|_F^2 \\ \text{s.t. } E\{\mathbf{H}\mathbf{H}^H\} &= \mathbf{U}_s (\mathbf{\Sigma}_s^2 - \sigma_N^2 \mathbf{I}) \mathbf{U}_s^H \end{aligned} \quad (15)$$

The detailed derivation to the (15) is provided in **Appendix A**, and the optimal solution is

$$\mathbf{B}_{opt} = \mathbf{R}_{\mathbf{H}} (\mathbf{R}_{\bar{\mathbf{H}}})^{-1}. \quad (16)$$

Combining (12), (13), (14), and (16), we obtain the enhanced CSI estimation as

$$\begin{aligned} \bar{\mathbf{H}}_u &= \mathbf{R}_{\mathbf{H}} (\mathbf{R}_{\bar{\mathbf{H}}})^{-1} \bar{\mathbf{H}} \\ &= \mathbf{U}_s (\mathbf{\Sigma}_s^2 - \sigma_N^2 \mathbf{I}) \mathbf{\Sigma}_s^{-2} \mathbf{U}_s^H \bar{\mathbf{H}} \\ &= \mathbf{U}_s (\mathbf{\Sigma}_s^2 - \sigma_N^2 \mathbf{I}) \mathbf{\Sigma}_s^{-2} \mathbf{U}_s^H \mathbf{U}_s \mathbf{\Sigma}_s \mathbf{V}_s^H, \\ &= \mathbf{U}_s (\mathbf{\Sigma}_s^2 - \sigma_N^2 \mathbf{I}) \mathbf{\Sigma}_s^{-1} \mathbf{V}_s^H \end{aligned} \quad (17)$$

where σ_N^2 can be replaced by $\hat{\sigma}_N^2$ to obtain the approximate value of $\bar{\mathbf{H}}_u$. Moreover, $\bar{\mathbf{H}}_u$ can be further expressed as $\bar{\mathbf{H}}_u = \mathbf{H} + \bar{\mathbf{N}}_u$, where $\bar{\mathbf{N}}_u$ is the suppressed noise.

Next, $\bar{\mathbf{H}}_u$ can be utilized for enhancing communication demodulation.

3) *Spatial Filtering*: In order to further estimate the range of UE for localization, we apply a baseband spatial filter for receiving the signals in $\hat{\mathbf{p}}_{R,0}^U = (\hat{\varphi}_0, \hat{\theta}_0)$. The BF vector for receiving the signals from the AoA of $\hat{\mathbf{p}}_{R,0}^U$ is generated with the LS method as [17]

$$\mathbf{w}_R = \frac{[\mathbf{a}^T(\hat{\mathbf{p}}_{R,0}^U)]^\dagger}{\sqrt{\|[\mathbf{a}^T(\hat{\mathbf{p}}_{R,0}^U)]^\dagger\|_2}} \in \mathbb{C}^{P_t Q_t \times 1}, \quad (18)$$

where $[\cdot]^\dagger$ represents the pseudo-inverse of a matrix.

Using \mathbf{w}_R to filter $\bar{\mathbf{H}}_u$, we obtain the filtered vector $\bar{\mathbf{h}}_R = (\mathbf{w}_R)^H \bar{\mathbf{H}}_u \in \mathbb{C}^{1 \times N_c M_s}$, and the $[(m-1)N_c + n]$ th column of $\bar{\mathbf{h}}_R$ can be expressed as

$$\begin{aligned} \bar{h}_{n,m} &= (\mathbf{w}_R)^H \mathbf{h}_{n,m} + \bar{n}_{n,m} \\ &= \sqrt{P_t^U} b_{C,0} \chi_{T,0} \chi_{R,0} e^{j2\pi m T_s^p \tilde{f}_{d,0,m}} e^{-j2\pi n \Delta f \tilde{\tau}_{0,m}} + \bar{n}_{n,m}, \end{aligned} \quad (19)$$

where $\tilde{f}_{d,0,m} = f_{d,0} + \delta_f(m)$, $\tilde{\tau}_{0,m} = \tau_0 + \delta_\tau(m)$, and $\chi_{R,0} = (\mathbf{w}_R)^H \mathbf{a}(\mathbf{p}_{R,0}^U)$ is the directional receiving gain of the signals. Moreover, the equivalent interference-plus-noise term is

$$\begin{aligned} \bar{n}_{n,m} &= (\mathbf{w}_R)^H \bar{\mathbf{n}}_{n,m} + \sum_{k=1}^{K-1} \left[\begin{array}{l} e^{j2\pi m T_s^p \tilde{f}_{d,k,m}} e^{-j2\pi n \Delta f \tilde{\tau}_{k,m}} \\ \times \sqrt{P_t^U} b_{C,k} \chi_{T,k} \chi_{R,I,k} \end{array} \right], \end{aligned} \quad (20)$$

where $\chi_{R,I,k} = (\mathbf{w}_R)^H \mathbf{a}(\mathbf{p}_{R,k}^U)$ ($k \neq 0$) is the receive gain of interference from the k th weak NLoS path, and $\bar{\mathbf{n}}_{n,m} = [\bar{\mathbf{N}}_u]_{:, (m-1)N_c + n}$. It is easy to see that $\|\chi_{R,I,k}\|_2^2 \ll \|\chi_{R,0}\|_2^2$.

B. MUSIC-based Range Estimation

Reshape $\bar{\mathbf{h}}_R$ to form a matrix $\bar{\mathbf{H}}_R \in \mathbb{C}^{N_c \times M_s}$ for range estimation, where $[\bar{\mathbf{H}}_R]_{n,m} = \bar{h}_{n,m}$. According to (19), $\bar{\mathbf{H}}_R$ has steering vector-like expressions, i.e., $e^{-j2\pi n \Delta f \tilde{\tau}_{k,m}}$ and $e^{j2\pi m T_s^p \tilde{f}_{d,k,m}}$, we can construct the range and Doppler steering vectors, respectively, as

$$\mathbf{a}_{r,m} = [e^{-j2\pi n \Delta f \frac{\tilde{\tau}_m}{c}}]_{n=0,1,\dots,N_c-1} \in \mathbb{C}^{N_c \times 1}, \quad (21)$$

$$\mathbf{a}_f = [e^{j2\pi m T_s^p \tilde{f}_{d,0,m}}]_{m=0,1,\dots,M_s-1} \in \mathbb{C}^{M_s \times 1}, \quad (22)$$

where $\tilde{\tau}_m = \tilde{\tau}_{0,m} \times c$ is the range of UE plus the offset due to TO.

Then, based on (19), the m th column of $\bar{\mathbf{H}}_R$ can be expressed by (21) and (22) as

$$[\bar{\mathbf{H}}_R]_{:,m} = S_0 \mathbf{a}_{r,m} [\mathbf{a}_f]_m + [\bar{\mathbf{N}}_R]_{:,m}, \quad (23)$$

where $S_0 = \sqrt{P_t^U} b_{C,0} \chi_{R,0} \chi_{T,0}$, and $[\bar{\mathbf{N}}_R]_{n,m} = \bar{n}_{n,m}$. We can use the MUSIC-based range estimation method [18] to estimate the range of UE based on the autocorrelation of $\bar{\mathbf{H}}_R$, as given by

$$\begin{aligned} \mathbf{R}_r &= \frac{1}{M_s} \bar{\mathbf{H}}_R (\bar{\mathbf{H}}_R)^H \\ &= \frac{1}{M_s} \sum_{m=0}^{M_s-1} [\bar{\mathbf{H}}_R]_{:,m} ([\bar{\mathbf{H}}_R]_{:,m})^H. \end{aligned} \quad (24)$$

Here, we further prove an important feature of the MUSIC-based range estimation method which is beneficial to localization under clock asynchronism in **Proposition 1**.

Proposition 1. *The MUSIC-based range estimation method can suppress the noise-like phase-shift terms due to TO.*

Proof. The detailed proof is provided in **Appendix B**. \square

Next, we briefly introduce the steps of MUSIC-based range estimation method. By applying eigenvalue decomposition (EVD) to \mathbf{R}_r , we obtain

$$\mathbf{R}_r = \mathbf{U}_r \boldsymbol{\Sigma}_r (\mathbf{U}_r)^H, \quad (25)$$

where $\boldsymbol{\Sigma}_r$ is the eigenvalue diagonal matrix, and \mathbf{U}_r is the corresponding eigen matrix. The noise subspace of \mathbf{R}_r for estimating the range of UE is $\mathbf{U}_{rN} = [\mathbf{U}_r]_{:, 2:N_c}$. Based on \mathbf{U}_{rN} , we can obtain the range spectrum function as

$$f_r(r) = \mathbf{a}_r(r)^H \mathbf{U}_{rN} (\mathbf{U}_{rN})^H \mathbf{a}_r(r), \quad (26)$$

where $\mathbf{a}_r(r) = [e^{-j2\pi n \Delta f \frac{r}{c}}]_{n=0,1,\dots,N_c-1} \in \mathbb{C}^{N_c \times 1}$ is the range steering vector.

The minimal point of $f_r(r)$ corresponds to the range to be estimated, and **Algorithm 1** can be used to identify the minimal point of $f_r(r)$. Note that $f(\mathbf{p})$, $\nabla_{\mathbf{p}} f(\mathbf{p})$ and $\mathbf{H}_{\mathbf{p}}(\mathbf{p})$ in **Algorithm 1** are replaced by (26), $\frac{\partial f_r(r)}{\partial r}$, and $\frac{\partial^2 f_r(r)}{\partial^2 r}$ for range estimation, respectively. The estimated range of UE is denoted by \hat{r}_0 .

C. Localization of UE

Since \hat{r}_0 corresponds to the estimated AoA of UE, $\hat{\mathbf{p}}_{R,0}^U = (\hat{\varphi}_0, \hat{\theta}_0)$, as shown in (18) and (19), \hat{r}_0 and $\hat{\mathbf{p}}_{R,0}^U$ form the polar coordinate of UE. Therefore, the Cartesian coordinate of UE can be expressed as

$$\boldsymbol{\Omega}_0 = (\hat{r}_0 \sin \hat{\theta}_0 \cos \hat{\varphi}_0, \hat{r}_0 \sin \hat{\theta}_0 \sin \hat{\varphi}_0, \hat{r}_0 \cos \hat{\theta}_0). \quad (27)$$

IV. JOINT CSI AND DATA SIGNALS-BASED LOCALIZATION SCHEME

In this section, we first demodulate the communication data signals based on the enhanced CSI obtained in Section III-A. Then, we propose a joint CSI and data signals-based localization scheme which exceeds the sensing accuracy of the localization scheme based on merely CSI as presented in Section III.

A. Data Signal Demodulation

The $[(m-1)N_c + n]$ th column of $\bar{\mathbf{H}}_u$ is the enhanced CSI estimates of $\mathbf{h}_{n,m}$. We use $\bar{\mathbf{h}}_{n,m} = [\bar{\mathbf{H}}_u]_{:, (m-1)N_c + n}$ to demodulate the received data signals and estimate the transmit data symbols. By applying the low-complexity zero-forcing to conduct channel equalization to $\mathbf{y}_{n,m}^i$, we obtain

$$\mathbf{y}_{n,m}^i = (\bar{\mathbf{h}}_{n,m})^\dagger \mathbf{y}_{n,m}^i. \quad (28)$$

Then, decoding the transmit data symbols based on $\mathbf{y}_{n,m}^i$ using the maximum-likelihood (ML) criterion, the estimation of the transmit data symbol at the n th subcarrier of the i th OFDM symbol of the m th packet can be expressed as

$$\hat{d}_{n,m}^i = \arg \min_{d \in \Theta_{\text{QAM}}} |y_{n,m}^i - d|^2, \quad (29)$$

where Θ_{QAM} is the used quadrature amplitude modulation (QAM) constellation.

Next, using the decoded data symbols and the received data signals, we can obtain the data-based CSI (D-CSI) as

$$\begin{aligned} \hat{\mathbf{h}}_{n,m}^i &= \frac{\mathbf{y}_{n,m}^i}{\hat{d}_{n,m}^i} \\ &= \mathbf{h}_{n,m} \frac{\hat{d}_{n,m}^i}{d_{n,m}^i} + \hat{\mathbf{n}}_{n,m}^i \\ &= \mathbf{h}_{n,m} \frac{1}{1 + \psi_{n,m}^i} + \hat{\mathbf{n}}_{n,m}^i, \end{aligned} \quad (30)$$

where $\psi_{n,m}^i = \frac{e_{n,m}^i}{d_{n,m}^i}$, $e_{n,m}^i = \hat{d}_{n,m}^i - d_{n,m}^i$ is the symbol decoding error, and $\hat{\mathbf{h}}_{n,m}^i = \frac{\mathbf{n}_{n,m}^i}{d_{n,m}^i}$ is the transformed Gaussian noise. We can see that when the transmit data symbol is decoded without error, i.e., $\psi_{n,m}^i = 1$, the estimated D-CSI is equal to the CSI estimated by LS method.

Stack all the estimated D-CSI $\hat{\mathbf{h}}_{n,m}^i$ at N_c subcarriers of M_s symbols to form a matrix $\hat{\mathbf{H}}^d \in \mathbb{C}^{P_t Q_t \times P_s M_s N_c}$, where $[\hat{\mathbf{H}}^d]_{:, [(i-1)M_s N_c + (m-1)N_c + n]} = \hat{\mathbf{h}}_{n,m}^i$ is the $[(i-1)M_s N_c + (m-1)N_c + n]$ th column of $\hat{\mathbf{H}}^d$, i.e., the D-CSI estimation at the n th subcarrier of the i th OFDM symbol of the m th packet.

B. Joint CSI and Data Signals-based Localization

In this subsection, we first present the joint CSI and data signals-based AoA and range estimation. Then, we introduce the single-base localization of UE based on the estimated AoA and range of UE.

1) *Joint CSI and Data Signals-based AoA Estimation:* By stacking all the D-CSI and enhanced P-CSI into one matrix, we obtain

$$\tilde{\mathbf{H}} = [\tilde{\mathbf{H}}_u, \hat{\mathbf{H}}^d] \in \mathbb{C}^{P_t Q_t \times (P_s+1)M_s N_c}, \quad (31)$$

where the first $M_s N_c$ columns of $\tilde{\mathbf{H}}$ is $\tilde{\mathbf{H}}_u$. According to (24), we can see that the MUSIC-based method does not require the input signals to be consecutive in subcarriers and OFDM symbols. Therefore, the MUSIC-based method can still be used to estimate the AoA and range of UE based on the autocorrelation of $\tilde{\mathbf{H}}$.

The autocorrelation of $\tilde{\mathbf{H}}$ can be expressed as

$$\begin{aligned} \mathbf{R}_{\tilde{\mathbf{H}}} &= \frac{1}{(P_s+1)M_s N_c} \tilde{\mathbf{H}} \tilde{\mathbf{H}}^H \\ &= \frac{1}{(P_s+1)M_s N_c} [\tilde{\mathbf{H}}_u (\tilde{\mathbf{H}}_u)^H + \hat{\mathbf{H}}^d (\hat{\mathbf{H}}^d)^H] \\ &= \frac{1}{(P_s+1)M_s N_c} [\tilde{\mathbf{H}}_u (\tilde{\mathbf{H}}_u)^H + \sum_{n,m,i} \hat{\mathbf{h}}_{n,m}^i (\hat{\mathbf{h}}_{n,m}^i)^H]. \end{aligned} \quad (32)$$

Then, we prove an important feature of $\mathbf{R}_{\tilde{\mathbf{H}}}$ that can improve the AoA sensing performance in **Proposition 2**.

Proposition 2. $\mathbf{R}_{\tilde{\mathbf{H}}}$ can be used in MUSIC-based method to enhance the coherent energy for AoA estimation.

Proof. The detailed proof is provided in **Appendix C**. \square

Subsequently, we present the MUSIC-based AoA estimation based on the joint CSI and data signals. By applying EVD to $\mathbf{R}_{\tilde{\mathbf{H}}}$, we obtain

$$\mathbf{R}_{\tilde{\mathbf{H}}} = \tilde{\mathbf{U}} \tilde{\Sigma} \tilde{\mathbf{U}}^H. \quad (33)$$

Then, the noise subspace obtained from $\mathbf{R}_{\tilde{\mathbf{H}}}$ is $\tilde{\mathbf{U}}_0 = [\tilde{\mathbf{U}}]_{:, 2:N_A}$, and the angle spectrum function can be expressed as

$$\tilde{f}_a(\mathbf{p}) = \mathbf{a}^H(\mathbf{p}) \tilde{\mathbf{U}}_0 (\tilde{\mathbf{U}}_0)^H \mathbf{a}(\mathbf{p}). \quad (34)$$

Subsequently, we use **Algorithm 1** to obtain the minimum points of $\tilde{f}_a(\mathbf{p})$ to estimate the AoA of UE. To identify the minimum of $\tilde{f}_a(\mathbf{p})$, we substitute $f(\mathbf{p})$, $\mathbf{H}_p(\mathbf{p})$, and $\nabla_p f(\mathbf{p})$ in **Algorithm 1** with (34), Hessian matrix, and the gradient vector of $\tilde{f}_a(\mathbf{p})$, respectively. Similar to the AoA estimation with pure CSI, the estimated angle with the largest $\tilde{f}_a^{-1}(\mathbf{p})$ is the AoA estimate of UE, which is denoted by $\tilde{\mathbf{p}}_{R,0}^U = (\tilde{\varphi}_0, \tilde{\theta}_0)$.

2) *Joint CSI and Data Signals-based Range Estimation:* Applying a baseband spatial filter to aggregate the signals from the direction of $\tilde{\mathbf{p}}_{R,0}^U$ for estimating the range of UE, we obtain

$$\tilde{\mathbf{h}}_R = (\tilde{\mathbf{w}}_R)^H \tilde{\mathbf{H}} \in \mathbb{C}^{1 \times (P_s+1)M_s N_c}, \quad (35)$$

where $\tilde{\mathbf{w}}_R = \frac{[\mathbf{a}^T(\tilde{\mathbf{p}}_{R,0}^U)]^\dagger}{\sqrt{\|[\mathbf{a}^T(\tilde{\mathbf{p}}_{R,0}^U)]^\dagger\|_2}} \in \mathbb{C}^{P_t Q_t \times 1}$ is the BF vector pointing at $\tilde{\mathbf{p}}_{R,0}^U$. Moreover, $\tilde{\mathbf{h}}_R$ can be expressed in a block matrix $\tilde{\mathbf{h}}_R = [\tilde{\mathbf{h}}_R, \tilde{\mathbf{h}}_R^d]$, where $\tilde{\mathbf{h}}_R^d = (\tilde{\mathbf{w}}_R)^H \hat{\mathbf{H}}^d \in \mathbb{C}^{1 \times P_s M_s N_c}$. Reforming $\tilde{\mathbf{h}}_R$ into a new block matrix for range estimation, we obtain

$$\tilde{\mathbf{H}}_R = [\tilde{\mathbf{H}}_R, \tilde{\mathbf{H}}_{R,1}^d, \dots, \tilde{\mathbf{H}}_{R,P_s}^d] \in \mathbb{C}^{N_c \times (P_s+1)M_s}, \quad (36)$$

where $\tilde{\mathbf{H}}_{R,i}^d \in \mathbb{C}^{N_c \times M_s}$ is the matrix stacked by the D-CSI estimation of the i th OFDM symbol in each packet. Moreover, the (n, m) th element of $\tilde{\mathbf{H}}_{R,i}^d$ is

$$\begin{aligned} [\tilde{\mathbf{H}}_{R,i}^d]_{n,m} &= (\tilde{\mathbf{w}}_R)^H \mathbf{h}_{n,m} \frac{1}{1+\psi_{n,m}^i} + \tilde{n}_{n,m}^i \\ &= \left(\frac{1}{1+\psi_{n,m}^i} \sqrt{P_t} b_{C,0} \chi_{T,0} \tilde{\chi}_{R,0} \right. \\ &\quad \left. \times e^{j2\pi m T_s^p f_{d,0,m}} e^{-j2\pi n \Delta f \tilde{\tau}_{0,m}} \right) + \tilde{n}_{n,m}^i, \end{aligned} \quad (37)$$

where $\tilde{\chi}_{R,0} = (\tilde{\mathbf{w}}_R)^H \mathbf{a}(\tilde{\mathbf{p}}_{R,0}^U)$, and $\tilde{n}_{n,m}^i$ is the transformed Gaussian noise with zero mean and variance σ_N^2 .

Since $\tilde{\mathbf{H}}_{R,i}^d$ also has the same range steering vector form as (21), we can express the m th column of $\tilde{\mathbf{H}}_{R,i}^d$ as

$$\tilde{\mathbf{h}}_{R,i,m}^d = [\tilde{\mathbf{H}}_{R,i}^d]_{:,m} = \tilde{S}_0 \tilde{\mathbf{a}}_{r,m}^i [\mathbf{a}_f]_m + [\tilde{\mathbf{N}}_{R,i}^d]_{:,m}, \quad (38)$$

where $\tilde{S}_0 = \sqrt{P_t} b_{C,0} \chi_{T,0} \tilde{\chi}_{R,0}$, and $\tilde{\mathbf{N}}_{R,i}^d$ is the transformed noise matrix with $[\tilde{\mathbf{N}}_{R,i}^d]_{n,m} = \tilde{n}_{n,m}^i$. Moreover, $\tilde{\mathbf{a}}_{r,m}^i$ is the range steering vector that can be distorted by the bit errors, which can be expressed as

$$\tilde{\mathbf{a}}_{r,m}^i = \left[\frac{1}{1+\psi_{n,m}^i} e^{-j2\pi n \Delta f \frac{\tilde{\tau}_m}{c}} \right]_{n=0,1,\dots,N_c-1} \in \mathbb{C}^{N_c \times 1}, \quad (39)$$

The MUSIC-based range estimation method exploits the autocorrelation of the signal matrix. The autocorrelation of $\tilde{\mathbf{H}}_R$ is expressed as

$$\begin{aligned} \tilde{\mathbf{R}}_r &= \frac{1}{(P_s+1)M_s} \tilde{\mathbf{H}}_R (\tilde{\mathbf{H}}_R)^H \\ &= \frac{1}{(P_s+1)M_s} \left[\tilde{\mathbf{H}}_R (\tilde{\mathbf{H}}_R)^H + \sum_{i=1}^{P_s} \sum_{m=0}^{M_s-1} \tilde{\mathbf{h}}_{R,i,m}^d (\tilde{\mathbf{h}}_{R,i,m}^d)^H \right]. \end{aligned} \quad (40)$$

Since $\tilde{\mathbf{H}}_R (\tilde{\mathbf{H}}_R)^H$ is the autocorrelation of pure CSI and it can definitely contribute to the coherent signal processing, the key issue is analyzing the coherence of

$\frac{1}{(P_s+1)M_s} \sum_{i=1}^{P_s} \sum_{m=0}^{M_s-1} \tilde{\mathbf{h}}_{R,i,m}^d (\tilde{\mathbf{h}}_{R,i,m}^d)^H$, which can be further expressed as

$$\begin{aligned} & \frac{1}{(P_s+1)M_s} \sum_{i=1}^{P_s} \sum_{m=0}^{M_s-1} \tilde{\mathbf{h}}_{R,i,m}^d (\tilde{\mathbf{h}}_{R,i,m}^d)^H \\ & \approx E_{i,m} \left\{ \tilde{\mathbf{h}}_{R,i,m}^d (\tilde{\mathbf{h}}_{R,i,m}^d)^H \right\} \\ & = E_{i,m} \left\{ \begin{aligned} & \left(\tilde{S}_0 \tilde{\mathbf{a}}_{r,m}^i [\mathbf{a}_f]_m + \tilde{\mathbf{n}}_{R,i,m}^d \right) \\ & \times \left(\tilde{S}_0 \tilde{\mathbf{a}}_{r,m}^i [\mathbf{a}_f]_m + \tilde{\mathbf{n}}_{R,i,m}^d \right)^H \end{aligned} \right\} \quad (41) \\ & = E_{i,m} \left\{ \left| \tilde{S}_0 \right|^2 \tilde{\mathbf{a}}_{r,m}^i (\tilde{\mathbf{a}}_{r,m}^i)^H + \tilde{\mathbf{n}}_{R,i,m}^d (\tilde{\mathbf{n}}_{R,i,m}^d)^H \right\} \\ & = \left| \tilde{S}_0 \right|^2 E_{i,m} \left\{ \tilde{\mathbf{a}}_{r,m}^i (\tilde{\mathbf{a}}_{r,m}^i)^H \right\} + \sigma_N^2 \mathbf{I}, \end{aligned}$$

where $\tilde{\mathbf{n}}_{R,i,m}^d = \left[\tilde{\mathbf{N}}_{R,i}^d \right]_{i,m}$ is the transformed noise. Here, we focus on the autocorrelation of the distorted range steering vector as shown in (39), i.e., $\mathbf{R}_{\tilde{\mathbf{r}}} = E_{i,m} \left\{ \tilde{\mathbf{a}}_{r,m}^i (\tilde{\mathbf{a}}_{r,m}^i)^H \right\}$. The (n_1, n_2) th element of $\mathbf{R}_{\tilde{\mathbf{r}}}$ can be expressed as

$$\begin{aligned} [\mathbf{R}_{\tilde{\mathbf{r}}}]_{n_1, n_2} &= \\ & e^{-j2\pi n' \Delta f \tau_k} E_{i,m} \left\{ \frac{1}{(1+\psi_{n_1,m}^i)(1+\psi_{n_2,m}^i)^*} e^{-j2\pi n' \Delta f \delta_r(m)} \right\}, \quad (42) \end{aligned}$$

where $n' = n_1 - n_2$. Since $\psi_{n,m}^i$ is related to the symbol error, we can divide the symbol error situation of all the OFDM data symbols in P_s packets into two cases to analyze the coherence of (41).

- 1: When the bit error rate is low enough and few or no symbol errors happen in the m th packet, then $[\mathbf{R}_{\tilde{\mathbf{a}}}]_{n_1, n_2}$ is equal to (51).
- 2: When the bit error rate is high, and there is at least one symbol error in each pair of (n_1, n_2) th subcarriers in an OFDM symbol.

The MUSIC-based range estimation method can effectively suppress the noise-like TO-related phase shift effectively based on the D-CSI estimation for situation 1 as shown in Section III-B.

Then, we focus on the situation 2 to analyze the influence of bit error rate (BER) on the coherence of signal autocorrelation for the MUSIC-based range estimation. In **Appendix D**, we prove that $[\mathbf{R}_{\tilde{\mathbf{r}}}]_{n_1, n_2} = 0$ in situation 2. Therefore, BER will not deteriorate the coherence and sensing accuracy of the MUSIC-based range estimation, which will also be shown in Section V. Using the MUSIC-based range estimation method in Section III-B by replacing \mathbf{R}_r with $\mathbf{R}_{\tilde{\mathbf{r}}}$, we can obtain the range estimation of UE, denoted by \tilde{r}_0 .

3) *Joint CSI and Data Signals-based Localization*: Based on the estimated AoA and range of UE obtained by the Joint CSI and data signal-based sensing scheme, we can further estimate the location of UE as

$$\tilde{\mathbf{\Omega}}_0 = (\tilde{r}_0 \sin \tilde{\theta}_0 \cos \tilde{\varphi}_0, \tilde{r}_0 \sin \tilde{\theta}_0 \sin \tilde{\varphi}_0, \tilde{r}_0 \cos \tilde{\theta}_0). \quad (43)$$

C. Complexity Analysis

The proposed joint single-base localization and communication enhancement scheme contains two MUSIC-based procedures to estimate AoA and range of UE, and a CSI

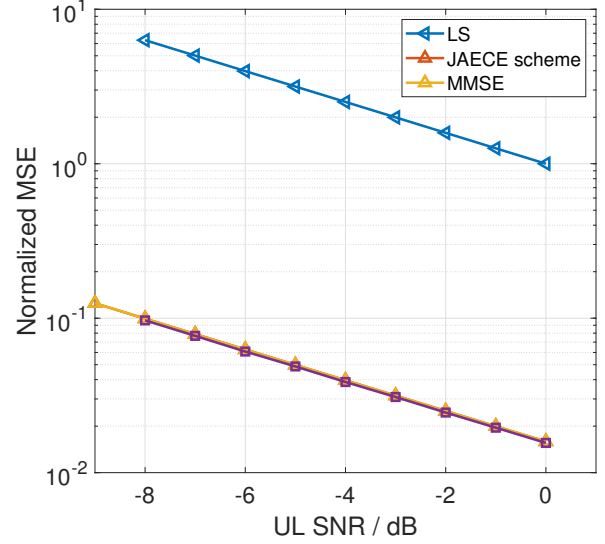


Fig. 3: The normalized MSE for CSI estimation.

enhancement procedure. Since the CSI enhancement procedure is deeply coupled with the AoA estimation procedure as shown in (17), the CSI enhancement procedure does not contribute extra complexity to the ISAC system. Therefore, the complexity of joint single-base localization and CSI enhancement in the preamble period as shown in Section III is $\mathcal{O}\{(P_t Q_t)^3 + N_c^3\}$.

The joint CSI and data signals-based localization adds $P_s M_s N_c$ D-CSI estimation based on OFDM data symbols to conduct MUSIC-based AoA and range estimation. Thus, the complexity of the joint CSI and data signal-based localization is approximately $\mathcal{O}\{(P_t Q_t)^3 + N_c^3 + (P_t Q_t)^2 (P_s + 1) M_s N_c + N_c^2 (P_s + 1) M_s\}$.

V. SIMULATION RESULTS

In this section, we present the communication and sensing simulation results of the proposed joint single-base localization and CSI enhancement scheme and the joint CSI and data signals-based localization scheme.

A. Simulation Setting

The carrier frequency is set to 28 GHz, the antenna interval, d_a , is half of the wavelength, the sizes of antenna arrays of the BS and user are $P_t \times Q_t = 8 \times 8$ and $P_r \times Q_r = 1 \times 1$, respectively. The subcarrier interval of UL preamble signal is $\Delta f = 480$ kHz, the subcarrier number is $N_c = 256$, and the bandwidth for JCAS is $B = N_c \Delta f = 122.88$ MHz. The variance of the Gaussian noise is $\sigma_N^2 = kFTB = 4.9177 \times 10^{-12}$ W, where $k = 1.38 \times 10^{-23}$ J/K is the Boltzmann constant, $F = 10$ is the noise factor, and $T = 290$ K is the standard temperature. The location of BS is (50, 4.75, 7) m. The location of UE is $(x_U, 0, 2)$ m, where x_U is uniformly distributed from 70 to 150 m. The velocity of UE is (-40, 0, 0) km/h, and the velocities of BS and the dumb scatterers are (0, 0, 0) m/s.

The range and location estimation MSEs are defined as the mean values of all the square errors of the range and location

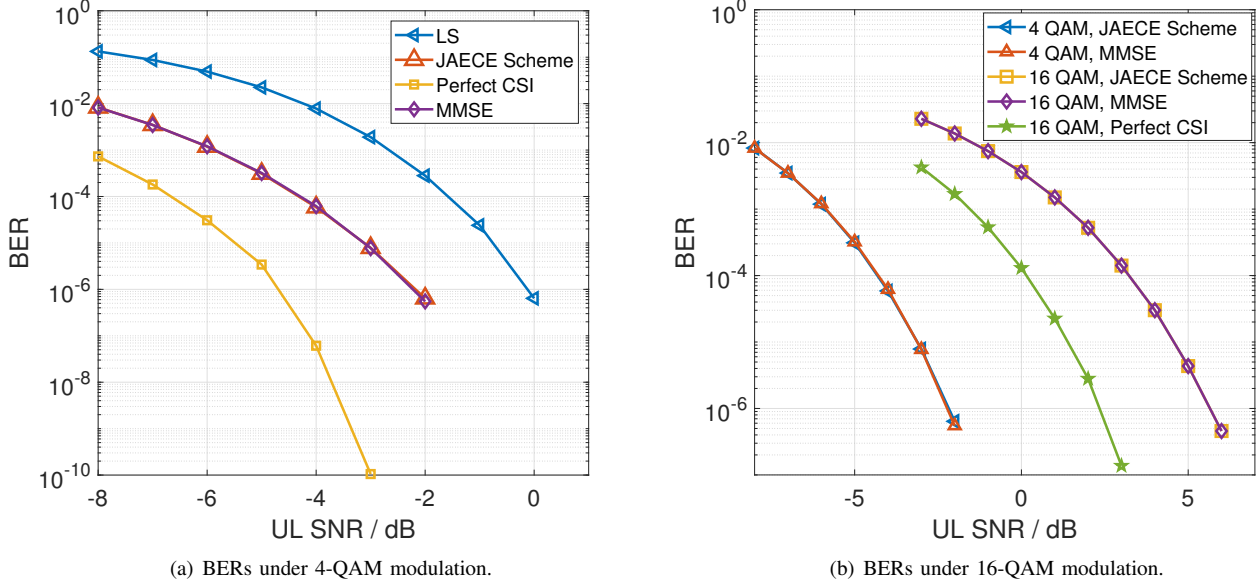
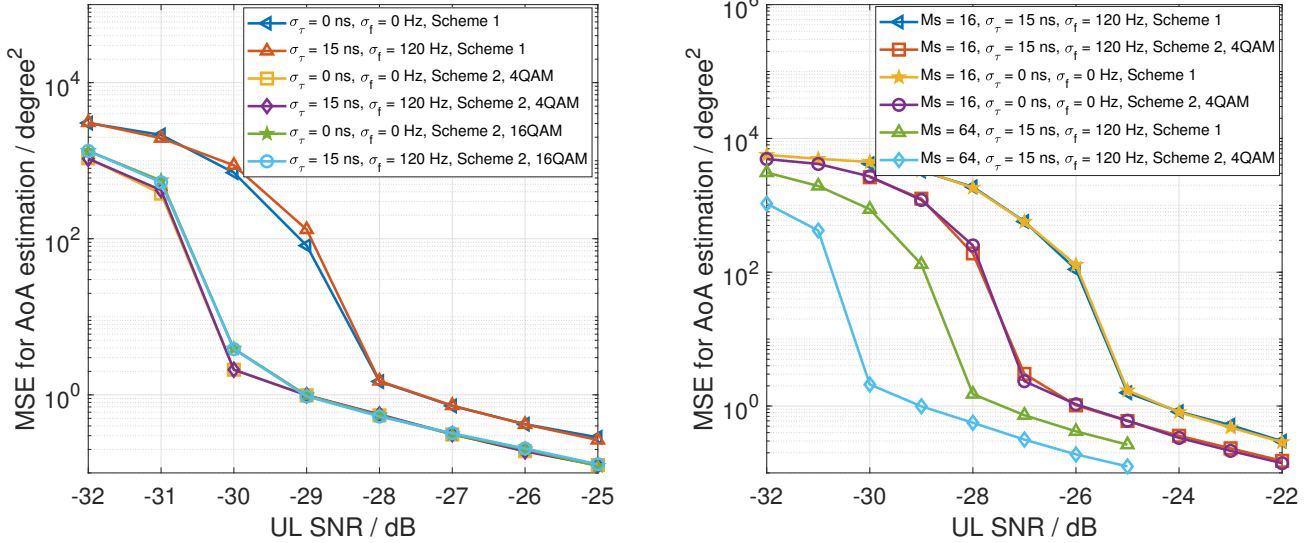


Fig. 4: The BERs using the JAECE scheme, LS, and MMSE methods under 4-QAM and 16-QAM modulation.



(a) AoA estimation MSEs of *Schemes 1* and *2* under various TOs, CFOs, and QAM orders when $M_s = 64$. (b) AoA estimation MSEs of *Schemes 1* and *2* under various TOs, CFOs, and QAM orders when $M_s = 64$.

Fig. 5: The AoA estimation MSEs of *Schemes 1* and *2* under various TOs, CFOs, M_s , and QAM orders.

estimation results under a certain set of simulation parameters. The RMSE is the square root of the MSE.

Based on the above locations and velocities of BS and user, the AoAs, AoDs, ranges, and Doppler shifts between UE and BS, and between the scatterers and BS can be derived to generate UL channel response matrix according to the models proposed in Section II-C. Further, BS can estimate the ranges and locations of UE according to Section III. Communication SNR is defined as the SNR of each antenna element of BS. According to (4), the UL communication SNR is expressed as

$$\gamma_c = \frac{P_t^U \sum_{k=0}^{K-1} |b_{C,k} \chi_{T,k}|^2}{\sigma_N^2}. \quad (44)$$

B. Communication Performance of the JAECE Scheme

This subsection presents the communication and AoA estimation performance of the proposed joint AoA estimation and CSI enhancement (JAECE) scheme in Section III. Moreover, we also show the AoA estimation performance of the joint CSI and data signal-based AoA estimation scheme proposed in Section IV-B.

Fig. 3 shows the CSI estimation Normalized MSEs (NMSEs) of the proposed JAECE scheme compared with the LS and MMSE CSI estimators. It can be seen that the CSI estimation NMSEs of the JAECE scheme approach those of the MMSE estimation method and are about 16 dB lower than those of the LS method.

Fig. 4 presents the BER performance of the enhanced CSI estimation of the proposed JAECE scheme compared with those using perfect CSI estimation and the CSIs estimated by the LS and MMSE methods under 4-QAM and 16-QAM modulation, respectively. Fig. 4(a) shows the BER performance under 4-QAM modulation. The BERs of demodulation using the JAECE scheme are generally the same as those using the MMSE method. Moreover, the JAECE scheme requires about 3 dB SNR lower than the LS method to achieve the same BER. Fig. 4(b) shows the BERs of demodulation using the proposed JAECE scheme and the MMSE method under 16-QAM modulation. It can be seen that the BERs of demodulation using the JAECE scheme still approach those using the MMSE method. It requires about 3 dB higher SNR for the JAECE scheme to achieve the same BER performance as the MMSE method.

C. Sensing Performance

We predefine *schemes* 1 and 2 for the simplicity of demonstration. *Scheme* 1 refers to the single-base localization scheme presented in Section III, while *scheme* 2 refers to the joint CSI and data signals-based localization scheme proposed in Section IV when $P_s = 1$.

Figs. 5 presents the AoA estimation MSEs of *schemes* 1 and 2 under various σ_τ , σ_f , M_s , and QAM orders. As SNR increases, the AoA estimation MSE decreases due to the accumulation of sensing energy. From Fig. 5(a), we can see that given the same M_s and QAM order, the AoA estimation MSEs of the same scheme are not affected prominently by the change of σ_τ and σ_f , which verifies the corresponding statement in Section III-A. The AoA estimation MSEs of *scheme* 2 are lower than those of *scheme* 1 for both situations when using 4 or 16 QAM modulation. This is because *scheme* 2 can accumulate the energy of both data signals and CSIs coherently to enhance the AoA estimation as proved in (32) in Section IV-B1. Moreover, given the same M_s , we can see that the AoA estimation MSEs of *scheme* 2 under 16 QAM are larger than those under 4 QAM. This is because the larger QAM order leads to larger equivalent noise as shown in (30). Fig. 5(b) shows the AoA estimation MSEs of *schemes* 1 and 2 under various σ_τ , σ_f and M_s . We can see that M_s influences the AoA estimation MSEs concretely. Given the same used scheme and QAM order, more OFDM packets used for AoA estimation results in lower AoA estimation MSEs. This is because the larger M_s contributes more coherent sensing energy as shown in (32).

Fig. 6 shows the MSEs of range estimation and localization of UE using *schemes* 1 and 2 under various σ_τ , σ_f and M_s when using 4-QAM modulation. The maximum-likelihood (ML)-based range estimation method [17] can be used to estimate the range and localize the targets when AoA is estimated. Figs. 6(a) and 6(b) present the range and location estimation MSEs of UE under various TOs and CFOs when $M_s = 64$. When $\sigma_\tau = 0$ ns and $\sigma_f = 0$ Hz, it refers to the ideal situation where TO and CFO do not exist.

Fig. 6(a) shows that when using the *scheme* 1, the MUSIC-based range estimation method can achieve about 6 dB lower

range estimation MSEs compared with the ML-based range estimation given the same σ_τ and σ_f . This verifies that the MUSIC-based method can suppress the TO-related noise-like terms exploiting the randomness of TO in various OFDM packets as shown in **Proposition 1**. Moreover, given the same σ_τ and σ_f , The range estimation MSEs of *scheme* 2 in the low SNR regime are lower than those of using *scheme* 1. This is because *scheme* 2 can accumulate more coherent energy for sensing as shown in (41).

Fig. 6(b) presents that the localization MSEs decrease as SNR increases since the AoA and range estimation MSEs both decrease as SNR increases according to Figs. 5 and 6(a). Given the same σ_τ and σ_f for *scheme* 1, we can see that the localization MSEs of using the MUSIC-based range estimation method are about 5 dB lower than those using the ML-based range estimation method. This is because the TO-related range ambiguity can be better suppressed than the conventional ML-based sensing method according to 6(a). Moreover, when *scheme* 2 and the MUSIC-based sensing method are used, the corresponding localization MSEs are about 8 dB lower than those with *scheme* 1 and the ML-based sensing method. This is because *scheme* 2 can further achieve better AoA estimation accuracy than *scheme* 1 as shown in Fig. 5. We can see that the combination of *scheme* 2 and the MUSIC-based sensing method can achieve decimeter-level single-base localization, while the conventional ML-based sensing method with *scheme* 1 can realize millimeter-level single-base localization.

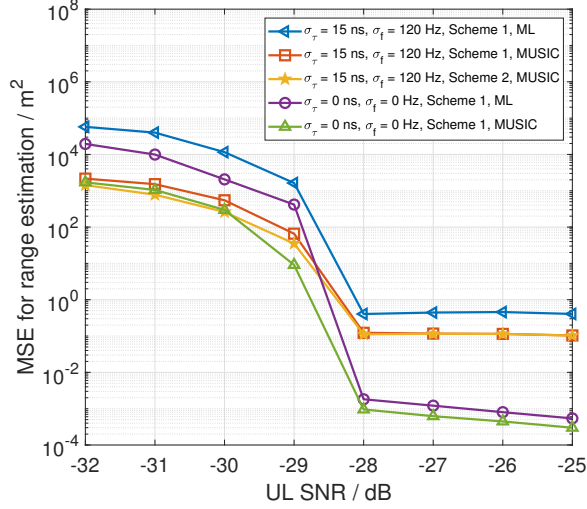
Figs. 6(c) and 6(d) show the range and location estimation MSEs of UE under various M_s when $\sigma_\tau = 15$ ns and $\sigma_f = 120$ Hz.

Fig. 6(c) shows that the larger M_s is, the smaller the MSEs of range estimation are. This is because the larger M_s contributes more coherent energy for range sensing as shown in (40) and (41). When $M_s = 64$, the required SNR to achieve the minimum range estimation MSE is about 3 dB lower than that for $M_s = 16$. Moreover, we can see that when M_s decreases, the range estimation MSE gap between the ML-based and MUSIC-based sensing methods becomes smaller. This is because more packets offer more terms to suppress the TO-related phase noise as shown in (51) in **Appendix B**.

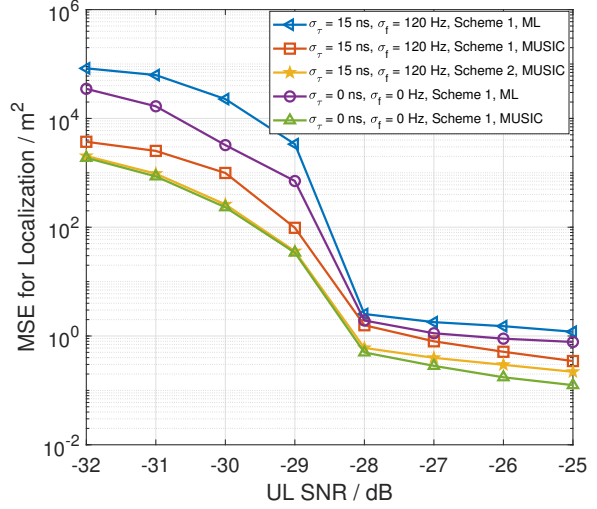
From Fig. 6(d), we can see that given the same scheme and the range estimation method, the smaller M_s leads to the larger localization MSEs, which results from the larger AoA and range estimation MSEs according to Figs. 5 and 6(c), respectively. Moreover, it is shown that the larger M_s is, the larger the localization MSE gap between using the *schemes* 1 and 2 is. This is because more packets offer more terms for suppressing the TO-related phase noise as shown in (51) in **Appendix B** and for accumulating the coherent energy for AoA estimation as shown in (32). Moreover, we can see that *scheme* 2 and the MUSIC-based sensing method can achieve decimeter-level single-base localization compared with the millimeter-level single-base localization achieved by the conventional ML-based sensing combined with *scheme* 1.

VI. CONCLUSION

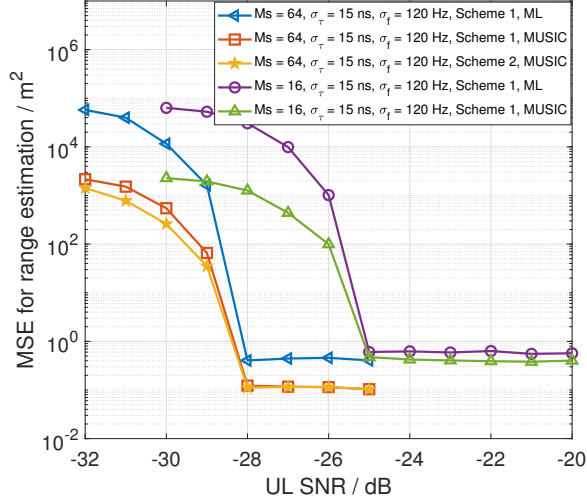
In this paper, we propose a joint single-base localization and communication enhancement scheme for the UL ISAC system.



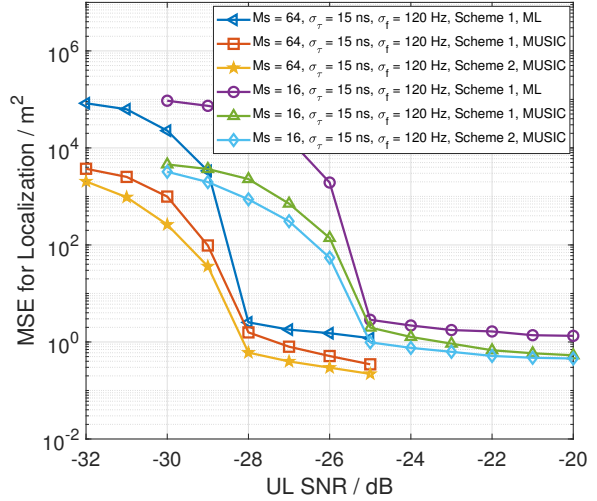
(a) The MSEs of range estimation under various TOs and CFOs when $M_s = 64$.



(b) The MSEs of localization under various TOs and CFOs when $M_s = 64$.



(c) The MSEs of range estimation under various M_s when $\sigma_\tau = 15$ ns and $\sigma_f = 120$ Hz.



(d) The MSEs of localization under various M_s when $\sigma_\tau = 15$ ns and $\sigma_f = 120$ Hz.

Fig. 6: The MSEs of range estimation and localization of *Schemes 1 and 2* under various TOs, CFOs, and M_s .

We first propose the JAECE scheme that integrates the CSI enhancement procedure into the MUSIC-based AoA estimation and imposes no additional complexity on the ISAC system. We further prove that the MUSIC-based range estimation method can suppress the time-varying TO-related phase terms by exploiting the averaging effects of the noise-like phase terms on the timeframe direction. Finally, we propose a joint CSI and data signals-based localization scheme that can coherently exploit the data signals with the CSI signals to improve the AoA and range estimation for localizing UE. Simulation results show that the BER performance of demodulation using the proposed JAECE scheme is equivalent to that using the MMSE method, and the localization MSEs using the proposed joint CSI and data signals-based localization scheme are about 8 dB lower than those of the ML-based benchmark method in the high SNR regime.

APPENDIX A SOLUTION TO (15)

Denote the objective of (15) as

$$\begin{aligned}
 J &= E \|\mathbf{B}\bar{\mathbf{H}} - \mathbf{H}\|_F^2 \\
 &= E \{ \text{Tr}[(\mathbf{B}\bar{\mathbf{H}} - \mathbf{H})(\mathbf{B}\bar{\mathbf{H}} - \mathbf{H})^H] \} \\
 &= E \{ \text{Tr}(\mathbf{B}\bar{\mathbf{H}}\bar{\mathbf{H}}^H\mathbf{B}^H + \mathbf{H}\mathbf{H}^H - \mathbf{B}\bar{\mathbf{H}}\mathbf{H}^H - \mathbf{H}\bar{\mathbf{H}}^H\mathbf{B}^H) \} \\
 &= \text{Tr}(\mathbf{B}\mathbf{R}_{\bar{\mathbf{H}}}\mathbf{B}^H + \mathbf{R}_{\mathbf{H}} - \mathbf{B}\mathbf{R}_{\mathbf{H}} - \mathbf{R}_{\mathbf{H}}\mathbf{B}^H)
 \end{aligned} \tag{45}$$

Since (45) is a convex problem, the optimal solution to (15), denoted by \mathbf{B}_{opt} , should satisfy

$$\left. \frac{\partial J}{\partial \mathbf{B}} = \mathbf{0} \right|_{\mathbf{B}=\mathbf{B}_{opt}}. \tag{46}$$

Combining (45) and (46), we obtain

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{B}} &= \mathbf{B} \left(\mathbf{R}_{\bar{\mathbf{H}}}^H + \mathbf{R}_{\bar{\mathbf{H}}} \right) - \mathbf{R}_{\mathbf{H}}^H - \mathbf{R}_{\mathbf{H}} \\ &= 2\mathbf{B}\mathbf{R}_{\bar{\mathbf{H}}} - 2\mathbf{R}_{\mathbf{H}} \end{aligned} \quad (47)$$

Finally, by combining (46) and (47), we derive the optimal solution to (15) as shown in (16).

APPENDIX B PROOF OF PROPOSITION 1

The autocorrelation of $\bar{\mathbf{H}}_R$ can be further expressed as

$$\begin{aligned} \mathbf{R}_r &= \frac{1}{M_s} \sum_{m=0}^{M_s-1} [\bar{\mathbf{H}}_R]_{:,m} \left([\bar{\mathbf{H}}_R]_{:,m} \right)^H \\ &= \frac{1}{M_s} \sum_{m=0}^{M_s-1} \left[|S_0[\mathbf{a}_f]_m|^2 \mathbf{a}_{r,m} (\mathbf{a}_{r,m})^H + \bar{\mathbf{n}}_{u,m} (\bar{\mathbf{n}}_{u,m})^H \right. \\ &\quad \left. + 2\text{Re} \left(S_0[\mathbf{a}_f]_m \mathbf{a}_{r,m} (\bar{\mathbf{n}}_{u,m})^H \right) \right], \end{aligned} \quad (48)$$

where $\bar{\mathbf{n}}_{u,m} = [\bar{\mathbf{N}}_R]_{:,m}$. Since $\bar{\mathbf{n}}_{u,m}$ is irrelevant to $\mathbf{a}_{r,m}$, \mathbf{R}_r can be further expressed as

$$\mathbf{R}_r = \frac{1}{M_s} \sum_{m=0}^{M_s-1} \left[|S_0[\mathbf{a}_f]_m|^2 \mathbf{a}_{r,m} (\mathbf{a}_{r,m})^H + \bar{\mathbf{n}}_{u,m} (\bar{\mathbf{n}}_{u,m})^H \right]. \quad (49)$$

We focus on the signal part of \mathbf{R}_r to analyze the suppression of TO-related phase-shift terms. The autocorrelation contributed by the useful signal can be expressed by

$$\mathbf{R}_{r,s} = \frac{|S_0|^2}{M_s} \sum_{m=0}^{M_s-1} \left[\mathbf{a}_{r,m} (\mathbf{a}_{r,m})^H \right]. \quad (50)$$

By substituting (21) into (50), the (n_1, n_2) th element of $\mathbf{R}_{r,s}$ can be expressed as

$$[\mathbf{R}_{r,s}]_{n_1, n_2} = \frac{|S_0|^2 e^{-j2\pi n' \Delta f \tau_k}}{M_s} \left(\sum_{m=0}^{M_s-1} e^{-j2\pi n' \Delta f \delta_\tau(m)} \right), \quad (51)$$

where $n' = n_1 - n_2$. Since $\delta_\tau(m)$ follows zero-mean stochastic distribution, $\delta_\tau(m)$ can be reduced to nanosecond-level [6], and Δf is from 15 kHz to 480 kHz, $\frac{1}{M_s} \left[\sum_{m=0}^{M_s-1} e^{-j2\pi n' \Delta f \delta_\tau(m)} \right]$ approaching 1 is satisfied for most n' when $M_s \rightarrow \infty$.

Therefore, the MUSIC-based range estimation method can suppress the TO-related phase shift.

APPENDIX C PROOF OF PROPOSITION 2

Since $\bar{\mathbf{H}}_u (\bar{\mathbf{H}}_u)^H$ is the P-CSI related term and will definitely contribute to the coherent energy for sensing, $\sum_{n,m,i} \hat{\mathbf{h}}_{n,m}^i (\hat{\mathbf{h}}_{n,m}^i)^H$ is the key part for analyzing whether the joint CSI and data signals can be used to improve the sensing performance. According to (30), we obtain

$$\begin{aligned} &\sum_{n,m,i} \hat{\mathbf{h}}_{n,m}^i (\hat{\mathbf{h}}_{n,m}^i)^H \\ &= \sum_{n,m,i} \left(\mathbf{h}_{n,m} \frac{1}{1+\psi_{n,m}^i} + \hat{\mathbf{n}}_{n,m}^i \right) \left(\mathbf{h}_{n,m} \frac{1}{1+\psi_{n,m}^i} + \hat{\mathbf{n}}_{n,m}^i \right)^H \\ &= \sum_{n,m,i} \left(\left| \frac{1}{1+\psi_{n,m}^i} \right|^2 \mathbf{h}_{n,m} (\mathbf{h}_{n,m})^H + \hat{\mathbf{n}}_{n,m}^i (\hat{\mathbf{n}}_{n,m}^i)^H \right). \end{aligned} \quad (52)$$

Note that in (52), we omit the cross-multiplication of $\hat{\mathbf{n}}_{n,m}^i$ and $\hat{\mathbf{h}}_{n,m}^i$ because its expectation is 0 in (32).

Since $\psi_{n,m}^i = \frac{e_{n,m}^i}{d_{n,m}^i}$ in (52) is not related to the antenna-relevant indexes, i.e., p and q , the sensing energy for AoA estimation can still be accumulated coherently. However, as the QAM-order used for modulation becomes higher, the BER increases and $\left| \frac{1}{1+\psi_{n,m}^i} \right|^2$ decreases, resulting in smaller processing SNR for AoA estimation.

APPENDIX D PROOF OF $[\mathbf{R}_{\hat{\mathbf{r}}}]_{n_1, n_2} = 0$ IN SITUATION 2

Here, $[\mathbf{R}_{\hat{\mathbf{r}}}]_{n_1, n_2}$ can be rewritten as

$$\begin{aligned} [\mathbf{R}_{\hat{\mathbf{r}}}]_{n_1, n_2} &= e^{-j2\pi n' \Delta f \tau_k} E_{i,m} \left\{ \frac{d_{n_1,m}^i (d_{n_2,m}^i)^*}{d_{n_1,m}^i (d_{n_2,m}^i)^*} e^{-j2\pi n' \Delta f \delta_\tau(m)} \right\}, \end{aligned} \quad (53)$$

where $d_{n_1,m}^i$ and $d_{n_2,m}^i$ are the actual data symbols, and $\hat{d}_{n_1,m}^i$ and $\hat{d}_{n_2,m}^i$ are the corresponding decoded symbols. Since the transmitted data symbols are irrelevant random symbols, and the symbol errors are caused by the i.i.d. Gaussian noise, we obtain $E_{i,m} \left\{ d_{n_1,m}^i (d_{n_2,m}^i)^* \right\} = 0$. Therefore, we can further derive

$$\begin{aligned} [\mathbf{R}_{\hat{\mathbf{r}}}]_{n_1, n_2} &= e^{-j2\pi n' \Delta f \tau_k} E_{i,m} \left\{ d_{n_1,m}^i (d_{n_2,m}^i)^* \right\} E_{i,m} \left\{ \frac{e^{-j2\pi n' \Delta f \delta_\tau(m)}}{d_{n_1,m}^i (d_{n_2,m}^i)^*} \right\} \\ &= 0 \end{aligned} \quad (54)$$

REFERENCES

- [1] Z. Feng, Z. Wei, X. Chen, H. Yang, Q. Zhang, and P. Zhang, "Joint Communication, Sensing, and Computation Enabled 6G Intelligent Machine System," *IEEE Network*, vol. 35, no. 6, pp. 34–42, Nov. 2021.
- [2] F. Liu, C. Masouros, A. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Transactions on Communications*, June 2020.
- [3] X. Chen, Z. Feng, Z. Wei, P. Zhang, and X. Yuan, "Code-Division OFDM Joint Communication and Sensing System for 6G Machine-Type Communication," *IEEE Internet of Things Journal*, vol. 8, no. 15, pp. 12 093–12 105, Feb. 2021.
- [4] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, "An overview of signal processing techniques for joint communication and radar sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1295–1315, Sept. 2021.
- [5] X. Yuan, Z. Feng, J. A. Zhang, W. Ni, R. P. Liu, Z. Wei, and C. Xu, "Spatio-Temporal Power Optimization for MIMO Joint Communication and Radio Sensing Systems With Training Overhead," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 514–528, Jan. 2021.
- [6] J. A. Zhang, K. Wu, X. Huang, Y. J. Guo, D. Zhang, and R. W. Heath, "Integration of radar sensing into communications with asynchronous transceivers," *IEEE Communications Magazine*, pp. 1–7, Aug. 2022.
- [7] S. A. Hassani, B. van Liempd, A. Bourdoux, F. Horlin, and S. Pollin, "Joint in-band full-duplex communication and radar processing," *IEEE Systems Journal*, pp. 1–9, July 2021.
- [8] G. Kwon, A. Conti, H. Park, and M. Z. Win, "Joint communication and localization in millimeter wave networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1439–1454, 2021.
- [9] X. Chen, Z. Feng, J. Andrew Zhang, Z. Wei, X. Yuan, and P. Zhang, "Sensing-aided uplink channel estimation for joint communication and sensing," *IEEE Wireless Communications Letters*, vol. 12, no. 3, pp. 441–445, March 2023.
- [10] W. Yuan, N. Wu, B. Etlzinger, Y. Li, C. Yan, and L. Hanzo, "Expectation-maximization-based passive localization relying on asynchronous receivers: Centralized versus distributed implementations," *IEEE Transactions on Communications*, vol. 67, no. 1, pp. 668–681, Jan. 2019.

- [11] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2.0: Passive human tracking with a single wi-fi link," *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, p. 350–361, 2018.
- [12] Z. Ni, J. A. Zhang, X. Huang, K. Yang, and J. Yuan, "Uplink sensing in perceptive mobile networks with asynchronous transceivers," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1287–1300, Feb. 2021.
- [13] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "FarSense: Pushing the Range Limit of WiFi-Based Respiration Sensing with CSI Ratio of Two Antennas," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 3, Sept. 2019.
- [14] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, "Multisense: Enabling multi-person respiration sensing with commodity wifi," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 4, no. 3, Sept. 2020.
- [15] X. Li, J. Andrew Zhang, K. Wu, Y. Cui, and X. Jing, "CSI-Ratio-based Doppler Frequency Estimation in Integrated Sensing and Communications," *IEEE Sensors Journal*, vol. 22, no. 21, pp. 20 886–20 895, Sept. 2022.
- [16] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1236–1259, May 2011.
- [17] J. A. Zhang, X. Huang, Y. J. Guo, J. Yuan, and R. W. Heath, "Multibeam for joint communication and radar sensing using steerable analog antenna arrays," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 671–685, Jan. 2019.
- [18] X. Chen, Z. Feng, Z. Wei, X. Yuan, P. Zhang, J. Andrew Zhang, and H. Yang, "Multiple signal classification based joint communication and sensing system," *IEEE Transactions on Wireless Communications*, pp. 1–1, Feb. 2023.
- [19] X. Chen, Z. Feng, J. Andrew Zhang, Z. Wei, X. Yuan, and P. Zhang, "Sensing-aided uplink channel estimation for joint communication and sensing," *IEEE Wireless Communications Letters*, vol. 12, no. 3, pp. 441–445, March 2023.
- [20] X. Chen, Z. Feng, Z. Wei, F. Gao, and X. Yuan, "Performance of Joint Sensing-Communication Cooperative Sensing UAV Network," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 15 545–15 556, Dec. 2020.
- [21] W. H. T. Rodger E. Ziemer, *Principles of Communications*, 7th ed. Wiley, 2014.
- [22] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Wiley Publishing, 2010.
- [23] M. Haardt, M. Pesavento, F. Roemer, and M. Nabil El Korso, "Chapter 15 - subspace methods and exploitation of special array structures," in *Academic Press Library in Signal Processing: Volume 3*, A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds. Elsevier, 2014, vol. 3, pp. 651–717.
- [24] B. De Moor, "The singular value decomposition and long and short spaces of noisy matrices," *IEEE Transactions on Signal Processing*, vol. 41, no. 9, pp. 2826–2838, Sept. 1993.