Distributed Multi-View Environment Sensing in Wireless Communication Networks

Citation for the original published paper (version of record):
Tong, X., Zhang, Z., Yang, Z. et al (2026). Distributed Multi-View Environment Sensing in Wireless Communication Networks. IEEE Transactions on Wireless Communications, 25: 12016-12033.
http://dx.doi.org/10.1109/TWC.2026.3659714

Distributed Multi-View Environment Sensing in Wireless Communication Networks

Xin Tong, *Member, IEEE,* Zhaoyang Zhang, *Senior Member, IEEE,* Zhaohui Yang, *Member, IEEE,*
Yu Ge, *Member, IEEE,* Henk Wymeersch, *Fellow, IEEE*

***Abstract*—In this paper, we propose a generic distributed multi-view sensing framework to deal with the three crucial challenges, i.e., physics-inspired signal propagation modeling, large-scale region of interest (ROI) pixelation, and distributed sensing architecture and algorithm design, in typical ISAC (integrated sensing and communication)-enabled wireless networks. First, we propose a unique pixel coordinate system with non-uniform and affine pixel division and mapping, which helps to resolve the occlusion effects between pixels and can be well tailored for common computational imaging technology. Second, inspired by Huygens' principle, we propose a novel integral-form signal propagation model that well captures the overall amplitude and phase effects of the pixels with arbitrary shapes and sizes, thus leading to a greatly reduced quantization error and making it suitable for large-scale scenario sensing. The proposed system model allows free pairing of transmitters and receivers in the network and makes full use of the multi-view observations, further resolving the occlusion problem any individual receiver may encounter in large-scale environment sensing. We then propose a low computational complexity alternating optimization (AO) algorithm, which iteratively performs sparse reconstruction and occlusion detection to obtain accurate sensing results. Finally, the performance of the proposed method is analyzed. Extensive numerical results verify the effectiveness of the proposed method.**

***Index Terms*—Integrated sensing and communication (ISAC), distributed sensing, occlusion effect, multi-view sensing.**

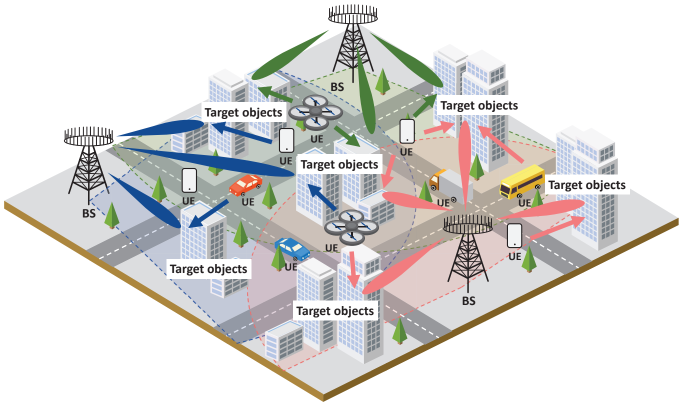


Fig. 1. The distributed multi-view ISAC system in the cellular wireless communication network.

I. INTRODUCTION

A. Motivation

IN the future sixth-generation (6G) wireless networks, many emerging applications like autonomous driving, unmanned aerial vehicles, intelligent robots and so on, require accurate environmental information including locations, shapes, and electromagnetic (EM) characteristics of the surrounding targets in large-scale scenarios [1]. One of the potential solutions is the emerging integrated sensing and communication (ISAC) technology [2], [3] which aims to utilize the ubiquitous wireless signals abound in the environment to achieve environment sensing within the wireless communication framework.

In the wireless communication scenario shown in Fig. 1, the complex environment will affect wireless channels. The first challenge in ISAC system design is how to solve a huge number of unknown environmental information, including the number, locations, and scattering characteristics of targets in the environment [4], [5]. Especially in large-scale scenarios, it is necessary to make full use of limited sensing resources, such as antennas and carriers, to design effective low-complexity ISAC methods. The trade-off between sensing accuracy and computational complexity is an essential basis for applying ISAC theory in reality.

The second challenge is how to deal with the complex propagation characteristics of EM waves and their impact on the targets and design a suitable propagation model [6]. We consider that a simple combination of scatter points cannot model the continuous shapes of targets in the environment. The continuous target will block the propagation

Xin Tong is with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, and with Zhejiang Provincial Laboratory of Multi-Modal Communication Networks and Intelligent Information Processing, Hangzhou 310027, China, and also with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden.(email: tongx@zju.edu.cn)

Zhaoyang Zhang is with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, and with the Institute of Fundamental and Transdisciplinary Research, Zhejiang University, Hangzhou 310058, China, and also with Zhejiang Provincial Laboratory of Multi-Modal Communication Networks and Intelligent Information Processing, Hangzhou 310027, China. (*Corresponding Author*, email: ning_ming@zju.edu.cn)

Zhaohui Yang is with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, and also with Zhejiang Provincial Laboratory of Multi-Modal Communication Networks and Intelligent Information Processing, Hangzhou 310027, China.(email: yang_zhaohui@zju.edu.cn)

Yu Ge and Henk Wymeersch are with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden. (emails: {yuge, henkw}@chalmers.se)

This work was supported in part by National Natural Science Foundation of China under Grants 62394292 and U20A20158, Zhejiang Provincial Key R&D Program under Grant 2023C01021, Ministry of Industry and Information Technology under Grant TC220H07E, and the Fundamental Research Funds for the Central Universities under Grant No. 226-2024-00069, and the SNS JU project 6G-DISAC under the EU's Horizon Europe research and innovation Program under Grant Agreement No 101139130.

of electromagnetic signals, preventing them from propagating further and making some receivers unable to receive these signals, which is called the *occlusion effect* [7], [8]. Therefore, each transceiver faces a different environment and has a unique measurement view, and multiple views from multiple transceivers should be exploited together. The joint processing of these unique views of an unknown environment also brings much difficulty. In addition, the reflection, scattering, and diffraction of EM waves by continuous objects should be considered, which are more complex than those of ideal point objects. Generally, the Huygens' principle can accurately describe the propagation characteristics of electromagnetic waves, but such a complex model not only makes the inversion of environmental information challenging, but is also unnecessary for communication. Therefore, the designed model needs to take into account both accuracy and complexity.

The last challenge to consider is to design a distributed multi-view sensing architecture in large-scale wireless communication scenarios [9], [10]. It would be best if jointly multi-view sensing of a large number of transceivers in a large-scale scenario could be achieved, but due to the limitations of sensing range and computational complexity, only partial joint sensing may be achieved. A large-scale scenario will be divided into multiple smaller local areas for sensing. At the same time, as the environment and the location of the transceiver change, the transceiver measuring each part will change. A general sensing architecture needs to enable any combination of transceivers in a large scenario to perform distributed environment sensing.

Motivated by the above three challenges, in this paper, we propose a low-complexity general distributed multi-view environment sensing architecture as an innovative ISAC system design in large-scale wireless communication scenarios.

B. Related Works

So far, some research works have proposed the idea of multi-view ISAC. From the perspective of multi-view architectures, in [7], the author describes the blueprint for multiple base stations (BSs) cooperative sensing in 6G. A framework of multi-BS cooperative sensing is proposed, which breaks through the limitation of single-BS sensing to achieve long-range and accurate sensing. In [10], both centralized and distributed architectures of multi-view sensing for wireless communications are proposed. For each architecture, the key performance indicators are pointed out.

From the perspective of multi-view sensing algorithms, [11] proposed a multiple ISAC devices outputs fusion method to achieve a higher sensing performance by exploiting multi-view data redundancy. A downlink mutual interference model of adjacent ISAC BSs was designed in [12] to deal with the downlink mutual interference, including sensing-related interference. In [13], the author proposed a vertical federated edge learning system for collaborative objects/human motion recognition by exploiting the multi-view wireless sensing data collected by distributed ISAC devices. In multistatic radio imaging, [14] and [15] achieve accurate sensing resolution and robustness through the collaboration of distributed nodes. In addition, there are still many potential environmental sensing technologies that can be applied to multi-view ISAC, multiple radar technology [16] , simultaneous localization and mapping (SLAM) [17] , distributed compressed sensing [18] , multi-view machine learning [19] , etc.

In addition to the above multi-view ISAC design ideas, computational imaging [20] as a potential environment sensing technology has recently been applied to wireless communication systems. Fully exploiting the sparse distribution characteristic of target objects in the environment, computational imaging is a high-resolution imaging method and has great potential to become an ISAC system design method. Different from conventional imaging methods, such as the radar method, the computational imaging method can obtain environmental information through pixel division, that is, dividing the unknown environment into discrete pixels as the smallest imaging unit and imaging by calculating the value of each pixel. Based on the compressed sensing (CS) theory [21] , the environment sensing problem has been transformed into the sparse reconstruction problem, which is solved by orthogonal matching pursuit (OMP) [22], generalized approximate message passing (GAMP) [23] and other widely used methods. Computational imaging has the advantage of enabling imaging with high resolution, especially for tiny targets. Consequently, the computational imaging method has many applications, including medical examinations and security checks. However, conventional computational imaging is usually limited to a small range, which cannot be directly applied to large-scale wireless communication scenarios. In large-scale wireless communication scenarios, high-resolution pixels usually result in a huge amount of calculation. Moreover, directly applying large pixels will inevitably lead to a huge quantization error caused by pixel division [4], [8]. The EM characteristics of the pixel point cannot represent the EM characteristics of the area where the pixel is located. Therefore, how to set the size of pixels to achieve a trade-off between calculation amount and sensing accuracy has become an unprecedented problem faced by computational imaging in wireless communication scenarios.

Some works have considered applying computational imaging to ISAC system design. Based on computational imaging and non-orthogonal multiple access technology, a millimeter wave ISAC system exploiting the sparsity of the environment is proposed in [4]. Considering the complex propagation characteristics of wireless signals, occlusion, diffraction, and diffraction characteristics are jointly considered in computational imaging technology to achieve multi-view sensing in wireless communication systems [8]. The above methods have proposed effective ISAC system design ideas, but they need to consider the errors caused by pixel division in computational imaging. In addition, the centralized processing architecture makes these methods difficult to apply in large-scale scenarios.

C. Main Ideas and Contributions

In this paper, we propose a novel general distributed multi-view sensing architecture with canceling the pixel division error, which extends computational imaging to large-scale scenario for the first time. Our design is depicted as follows. For architecture design, in the outdoor large-scale wireless communication scenario, we establish a pixel coordinate system for each BS and user equipment (UE) to represent the measurement view of its location. We divide the non-uniform pixel coordinates according to distance and angle, which intuitively describes the occlusion effect and the received power of pixels. Inspired by the application of Huygens' principle in antenna theory, we consider the continuous characteristics of object surfaces in the environment and propose a signal propagation model based on the integral form to achieve the cancellation of the phase error that plays an important role in large-size pixel division errors. Based on the overlap and position relationship between pixels, we propose a coordinate mapping method so that the pixel coordinate systems of different views can be mapped to each other. The proposed method supports a distributed multi-view architecture. In large-scale deployments, BSs and UEs can flexibly cooperate according to practical conditions. For instance, subsets of BSs may exchange data to reconstruct partial environments, or all BSs may participate jointly when a global environment reconstruction is desired. This flexibility allows the algorithm to scale to different network sizes while avoiding the bottleneck of centralized processing.

For algorithm design, we propose a joint multi-view sensing algorithm for the cooperating BSs and UEs. According to the proposed model, we jointly process the signals of multiple transceivers and carriers and convert the environment sensing problem into a CS reconstruction problem based on the sparse characteristics of the environment. To deal with the differences in views of different transceivers, we propose an alternating optimization (AO) algorithm based on GAMP. The proposed AO algorithm iteratively performs sparse reconstruction and occlusion detection to obtain accurate environment sensing results. Additionally, due to the pre-divided non-uniform pixels and mapping method, our proposed AO algorithm achieves low computational complexity in the large-range scenario.

The main contributions of this paper are as follows:

- We propose a general distributed multi-view sensing architecture. Based on non-uniform pixel division and mapping, we enable free cooperation between BSs and UEs in the environment and fully use each transceiver's unique view to achieve large-scale environment sensing.
- Based on the continuous characteristics of object surfaces, we propose a novel signal propagation model based on the integral form to achieve the cancellation of the phase error, which extends computational imaging to a large wireless communication scenario for the first time.
- To achieve joint multi-view sensing between the cooperating BSs and UEs. We propose a low computational complexity AO algorithm that iteratively performs sparse reconstruction and occlusion detection to obtain accurate environment sensing results.
- The performance of the proposed pixel division method and multi-view sensing method is analyzed, with extensive numerical results validating their effectiveness.

The rest of this paper is organized as follows. Section II presents the environment setting and system model in the large-scale communication scenario. Section III proposes the multi-view environment sensing algorithm. In Section IV, we analyze the impact of the proposed pixel division method and multi-view sensing method on the system performance. Finally, Section V presents the numerical results, and Section VI concludes the paper.

Notation: Fonts a, $\boldsymbol{a}$, and $\mathbf{A}$ represent scalars, vectors, and matrices, respectively. $\mathbf{A}^{\mathsf{T}}$ and $\|\mathbf{A}\|_{\mathrm{F}}$ denote transpose and Frobenius norm of $\mathbf{A}$, respectively. $|\cdot|$ and $[\cdot]$ denote the modulus and the catenation of the matrix, respectively. $\odot$ represents the Hadamard product between two matrices. Finally, notation $\mathrm{diag}(\boldsymbol{a})$ represents a diagonal matrix with the entries of $\boldsymbol{a}$ on its main diagonal, and $\delta(\cdot)$ is the Dirac delta function.

II. ENVIRONMENT SETTING

A. Distributed Multi-View ISAC Scenario

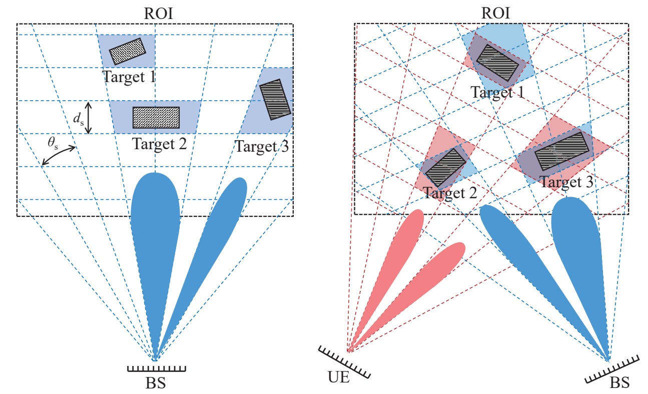
As shown in Fig. 1, we proposed a multi-view ISAC system in the cellular wireless communication network where multiple BSs and multiple active UEs such as cellphones, vehicles, and drones are deployed. In the cellular millimeter wave communication scenario, multiple UEs send uplink communication data to BSs. We consider that the transceivers are phase-level synchronized and UEs have been positioned with a small fraction of the wavelength.[1] A specific communication modulation method is adopted such as orthogonal frequency division multiplexing (OFDM). The transmitted signal is scattered by buildings and received by BSs through multipath channels. Therefore, the received signal of the BS contains environmental information, and the transceiver signals between the UEs and the BSs are exploited to detect and sense target objects such as buildings in the environment. Our purpose is to process and detect the environmental information in the communication signal to achieve environment sensing in such kind of wireless communication framework.

Addressing the occlusion problem is the research focus of the proposed distributed multi-view sensing scenario. As shown in Fig. 1, the BS can only receive signals from nearby target objects, while signals from farther away targets cannot be received because they are occluded or too far away. In addition, the occlusion effect will also occur on the front and back sides of the same target object, so that the BS can only receive signals from the side facing it. We need to jointly process multi-view signals from multiple BSs and UEs to deal with common occlusion effects and achieve complete sensing of the scenario. In addition, for distributed sensing, the fan-shaped area in Fig. 1 shows the sensing range of each BS. Each BS cooperates with UEs within a limited range to achieve partial sensing results in a large-scale environment.

As shown in Fig. 2, we consider a simplified 2-dimensional (2D) multi-view ISAC scenario, where multiple UEs act as

---

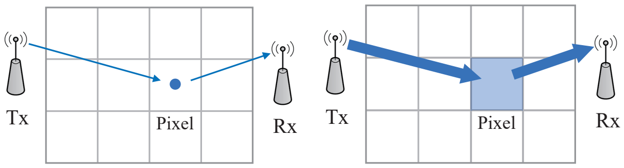
[1] We mainly focus on scenarios with stationary or quasi-stationary antennas, where phase-level synchronization and high-precision positioning are more feasible. For movable antennas, there exist emerging technologies that can enable high-accuracy positioning and facilitate synchronization [24]. In Section VI-C, we discuss mitigation strategies for practical deployments.



(a) Single pixel coordinate grid in a 2D scenario.

(b) Multiple pixel coordinate grids in a 2D scenario.

Fig. 2. The pixel division method in a 2D scenario.



(a) The conventional computational imaging model.

(b) The proposed model.

Fig. 3. Pixel propagation gain models in a 2D scenario.

transmitters (Txs) and send communication signals to multiple BS receivers (Rxs).[2] In the considered system model, we define the region of interest (ROI) as the environment to be sensed. The ROI represents the entire large-scale environment, and its location and size are determined by the sensing coverage of the coordinated BSs and UEs as well as the specific application requirements. For example, in an intelligent transportation scenario, the ROI can be set as the road area where vehicles and pedestrians are located, while in an industrial scenario, the ROI can be set as the interior space of a factory. The specific size and shape of the ROI do not affect the proposed model and algorithm, as they are designed to be irrelevant to the ROI geometry.

B. Pixel Division Method

In this section, we discretize the ROI using pixel-based representation, where each pixel characterizes the spatial distribution of objects and their electromagnetic scattering properties. In the conventional computational imaging method, the 2D environment is uniformly divided into multiple pixels. To achieve general distributed multi-view sensing, we discretize the environmental information within the ROI into non-uniform pixels and establish a unique affine coordinate system for each transceiver. Fig. 2(a) shows the pixels from the view of a single BS or UE. In the entire scenario, every UE and BS has a separate coordinate system. The method of dividing pixels in a 2D scenario is to divide the ROI into equal angle intervals according to the angle of departure/arrival (AOD/AOA) of the transceiver and to divide the ROI into several equal layers perpendicular to the direction of the array. In the 3D scenario, the pixels in the 2D ROI can be rotated into equal intervals to fill with pixels in the 3D scenario.

Let the interval between each divided angle be θ_s, and the interval between each layer is d_s. Therefore, the position of N_s pixels in the ROI is denoted as $\boldsymbol{p} \triangleq [\theta_\mathrm{s}, d_\mathrm{s}]$, where N_s denotes the number of pixels.[3] The proposed pixel division method intuitively describes the occlusion effect and the received power of pixels. In the same direction, targets located at closer pixels will occlude pixels farther away. For example, in Fig. 2(a), target 2 occludes target 1.

We use the scattering coefficient[4] x_{n_s} to represent the environmental information in the n_s-th pixel. If no target object in n_s-th pixel, $x_{n_\mathrm{s}} = 0$. Otherwise, we have $x_{n_\mathrm{s}} \in (0, 1]$. Therefore, the environmental information of the ROI can be characterized by the scattering coefficient vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_{N_\mathrm{s}}]^\mathsf{T}$. It is worth noting that $\boldsymbol{x}$ is a property of the environment itself. The environmental information only depends on the environment and is not related to the propagation path[5]. In the following sections, we add a subscript to $\boldsymbol{x}$ to indicate the chosen environment information vector of the coordinate systems under different transceiver views.

The targets within the ROI are located in different pixel grids for different transceivers. Fig. 2(b) shows the overlap between the two coordinate systems. As shown in Fig. 2(b), red and blue pixels are divided according to the positions of the UE and the BS, respectively, and the targets in the ROI are located in different pixels in the two coordinate grids. At the same time, in the proposed pixel division method, for a single transceiver, targets in pixels that are distributed closer in one direction will cause the occlusion effect of farther pixels, and the environment observed by different transceivers is different. For example, in the red grid from the view of the UE, target 2 occludes target 1. In the blue grid from the view of the BS, target 3 occludes target 1. Based on the different views of each transceiver, each view has a partial-view measurement equation. We jointly solve the measurement equations of multiple transceivers to achieve a complete multi-view environment sensing of the ROI. Meanwhile, since it is a pre-divided pixel grid, the occlusion relationship between pixels can be directly defined before the imaging process, which effectively reduces the computational complexity of the sensing algorithm.

III. SYSTEM MODEL

A. Propagation Model with Pixel Division Error Cancellation

The proposed environment sensing method based on computational imaging divides the environment into several pixels,

---

[2] It is noteworthy that the analysis method for 3-dimensional (3D) scenarios can be derived using a similar approach.

[3] We use quadrilateral pixels as shown in Fig. 2, which is different from polar coordinates. In addition, the proposed method can be performed with pixels of any shape.

[4] Diffraction is typically much weaker than the dominant scattering paths and is therefore neglected in the main model. When its contribution becomes non-negligible, it can be naturally absorbed into the estimated scattering coefficient. The pixel-based model can also be extended to explicitly include diffraction and other propagation effects [25].

[5] We assume that targets do not cause any phase rotation of the signal.

and the environment sensing result is obtained by calculating the scatter coefficients of pixels. In this section, we describe the proposed propagation model based on a global coordinate system. In the following sections, this global coordinate system can be replaced by an arbitrary coordinate system from the view of any transceiver. As mentioned in Section I-B, how to eliminate pixel division errors in the propagation model is the key to achieving large-scale environment sensing. The specific environment sensing method is as follows. To ensure high sensing accuracy, we consider multiple subcarriers. For the n_f-th subcarrier, the communication signal between transceiver pair propagates through multipath channels. The multipath propagation gain $\mathbf{H}_{n_\mathrm{f}}^{\mathrm{NLOS}} \in \mathbb{C}^{N_\mathrm{R}\times N_\mathrm{T}}$ is mainly composed of two parts[6], i.e. the free space propagation gain $\tilde{\mathbf{H}}_{n_\mathrm{f}}^{\mathrm{Tx}\to\mathrm{P}} \in \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{T}}$ from the Txs to pixels and the free space propagation gain $\tilde{\mathbf{H}}_{n_\mathrm{f}}^{\mathrm{P}\to\mathrm{Rx}} \in \mathbb{C}^{N_\mathrm{R}\times N_\mathrm{s}}$ from pixels to the Rxs, where N_T is the number of Txs and N_R is the number of Rxs, which are also the number of UEs and BSs in the scenario, respectively.

As shown in Fig. 3(a), take rectangular pixels as an example, the above-mentioned free space propagation gain is calculated through a statistical channel model based on the antenna position and pixel position in conventional computational imaging methods. Let the position of n_T-th Tx be $(x_{n_\mathrm{T}}, y_{n_\mathrm{T}})$ and let the position of n_s-th pixel center point be $(x_{n_\mathrm{s}}, y_{n_\mathrm{s}})$, the free space propagation gain from the n_T-th Tx to the n_s-th pixel is calculated as

$$h_{n_\mathrm{f}}^{\mathrm{Tx}n_\mathrm{T}\to\mathrm{P}n_\mathrm{s}} = \frac{\lambda_{n_\mathrm{f}}}{4\pi d} e^{-2j\pi d/\lambda_{n_\mathrm{f}}}, \tag{1}$$

where $d = \sqrt{(x_{n_\mathrm{T}} - x_{n_\mathrm{s}})^2 + (y_{n_\mathrm{T}} - y_{n_\mathrm{s}})^2}$, λ_{n_f} is the wavelength of the n_f-th subcarrier and d is the distance between the antenna and the pixel. In addition, $h_{n_\mathrm{f}}^{\mathrm{P}n_\mathrm{s}\to\mathrm{Tx}n_\mathrm{R}}$, $h_{n_\mathrm{f}}^{\mathrm{Rx}n_\mathrm{T}\to\mathrm{P}n_\mathrm{s}}$ and $h_{n_\mathrm{f}}^{\mathrm{P}n_\mathrm{s}\to\mathrm{Rx}n_\mathrm{R}}$ can be the computed in a similar way.

However, it is impractical to directly apply the conventional computational imaging propagation model as shown in (1) to environment sensing. The reason is that the phase errors caused by dividing into pixels have the greatest impact on computational imaging performance [4], [8]. In Fig. 3(a), when the target in the environment deviates from the center of the pixel, a phase error will occur. There is a wave path difference between the path from the antenna to the pixel and the path from the antenna to the point target, which causes the phase error in the propagation gain of the two paths. Therefore, when the pixel size exceeds one wavelength, the phase error becomes almost a random error, and this error is more obvious in planar targets. In conclusion, the propagation phase characteristic from the center point of the pixel to the antenna cannot represent the propagation phase characteristics of all locations within this pixel range. The specific theoretical analysis of errors is provided in Section V-A.

The conventional computational imaging method can only support tiny pixels, which results in extremely high sensing resource overhead, such as the number of antennas, number of carriers, storage space, computational complexity, etc. In order to achieve the sensing of the large-scale environment, we propose a method of using large pixels for computational imaging inspired by the application of Huygens' principle in antenna theory. According to the application of Huygens' principle in antenna theory, the antenna aperture surface is divided into many surface elements as secondary radiation sources, and the entire aperture surface radiation field is calculated by summing the radiation from all secondary radiation sources, which is also called Huygens sources. Based on the above theory, considering the continuous characteristics of the target surface, we also divide the considered scenario into pixels but use an integral method to precisely calculate the propagation gain from the antenna to the pixel position, as shown in Fig. 3(b). We calculate the impact of the entire pixel area on the propagation gain and the propagation model in (1) that can be reformulated similarly to the summation of Huygens sources. As shown in Fig. 3(b), part of propagation gain from the n_T-th Tx to the n_R-th Rx which scattered by the n_s-th pixel is calculated as

$$h_{n_\mathrm{f}}^{\mathrm{Tx}n_\mathrm{T}\to\mathrm{P}n_\mathrm{s}\to\mathrm{Rx}n_\mathrm{R}} = \iint_{\boldsymbol{S}_{n_\mathrm{s}}} \frac{h_{n_\mathrm{f}}^{\mathrm{Tx}n_\mathrm{T}\to\mathrm{P}n_\mathrm{s}} h_{n_\mathrm{f}}^{\mathrm{P}n_\mathrm{s}\to\mathrm{Rx}n_\mathrm{R}}}{S_{n_\mathrm{s}}} \mathrm{d}x\mathrm{d}y, \tag{2}$$

$$h_{n_\mathrm{f}}^{\mathrm{Tx}n_\mathrm{T}\to\mathrm{P}n_\mathrm{s}} = \frac{\lambda_{n_\mathrm{f}}}{4\pi d_\mathrm{T}} e^{\frac{-2j\pi d_\mathrm{T}}{\lambda_{n_\mathrm{f}}}}, \tag{3}$$

$$d_\mathrm{T} = \sqrt{(x_{n_\mathrm{T}} - x)^2 + (y_{n_\mathrm{T}} - y)^2}, \tag{4}$$

$$h_{n_\mathrm{f}}^{\mathrm{P}n_\mathrm{s}\to\mathrm{Rx}n_\mathrm{R}} = \frac{\lambda_{n_\mathrm{f}}}{4\pi d_\mathrm{R}} e^{\frac{-2j\pi d_\mathrm{R}}{\lambda_{n_\mathrm{f}}}}, \tag{5}$$

$$d_\mathrm{R} = \sqrt{(x_{n_\mathrm{R}} - x)^2 + (y_{n_\mathrm{R}} - y)^2}, \tag{6}$$

where $(x, y) \in \boldsymbol{S}_{n_\mathrm{s}}$, $\boldsymbol{S}_{n_\mathrm{s}}$ are the set of differential points within n_s-th pixel range, d_T and d_R is the distance from the differential point to the transmitting and receiving antennas, respectively. The area of n_s-th pixel is denoted by S_{n_s}. It is worth noting that we follow the cascade channel approach widely adopted in radar and wireless communications [27], [28] to obtain the product term $h_{n_\mathrm{f}}^{\mathrm{Tx}n_\mathrm{T}\to\mathrm{P}n_\mathrm{s}} h_{n_\mathrm{f}}^{\mathrm{P}n_\mathrm{s}\to\mathrm{Rx}n_\mathrm{R}}$ in (2). The propagation gain over the n_s-th irregularly shaped pixel is obtained by numerically integrating the propagation gain over its coverage area[7] $\boldsymbol{S}_{n_\mathrm{s}}$. In practice, $\boldsymbol{S}_{n_\mathrm{s}}$ is uniformly partitioned into sufficiently small subregions, and the integral is approximated by the summation of the gain contributions from all subregions. Since the pixel division is fixed once the transceivers are deployed, these integrals can be computed and stored in advance, thus avoiding repeated computations during the distributed execution of the proposed algorithm.

The proposed computational imaging propagation model can effectively cancel the effects of phase errors. If the pixels are divided on a uniform plane, then the error of pixel division can be completely avoided. Only when the pixel area exceeds the area of the target object, or the scattering coefficient of the target object is uneven, errors will be caused in the

---

[6] In this paper we assume that the existing LOS channel can be calculated by the free space propagation model and removed from the channel estimation result because the LOS channel does not contain unknown environmental information. Higher-order interactions are also ignored and treated as noise due to the large attenuation of the mmWave NLoS path [26].

[7] From the modeling perspective, (2) is general. Given the geometry and scattering properties of an object, the integration domain can be either a plane or a curved surface, and the scattering coefficient can also be spatially non-uniform.

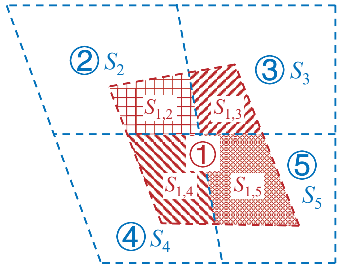


Fig. 4. The proposed coordinate mapping method.

pixel division process. However, these are errors in amplitude and have little impact on the sensing accuracy. The specific theoretical analysis of errors is given in Section V.

B. Coordinate Mapping

In order to jointly process data from multiple BS views and multiple UE views, we propose a coordinate mapping method to achieve mutual mapping of pixels from different views. The proposed method specifically includes two parts: power mapping and phase mapping.

a) Power Mapping: As shown in Fig. 4, we propose a power mapping method based on pixel area. For example, let the pixels from the view of BS A be red, and the pixels from the view of BS B be blue. Red pixel 1 overlaps with four blue pixels. We multiply the scattering coefficients of blue pixels 2, 3, 4, and 5 by different weights (power mapping coefficients) according to the size of the overlapping area $S_{1,2}, S_{1,3}, S_{1,4}, S_{1,5}$ to obtain the scattering coefficients, i.e.,

$$x_1^{\mathrm{BSA}} = \frac{S_{1,2}}{S_2} x_2^{\mathrm{BSB}} + \frac{S_{1,3}}{S_3} x_3^{\mathrm{BSB}} + \frac{S_{1,4}}{S_4} x_4^{\mathrm{BSB}} + \frac{S_{1,5}}{S_5} x_5^{\mathrm{BSB}}, \quad (7)$$

where each x represents the scattering coefficient of each pixel.

b) Phase Mapping: The phase mapping step essentially ensures pixel mappings synchronization after transceiver synchronization. We perform phase mapping based on the positional relationship between the pixel and the transceiver. We consider that the free space channel without the unknown environmental information can be modeled as a deterministic geometric channel. Based on the free space propagation gain calculated in (2), the phase mapping coefficient is calculated as the phase difference between the paths before and after mapping. For example, for the n_{f}-th subcarrier, the n_{T}-th UE and the n_{R}-th BS, the phase mapping coefficient mapping the $n_{\mathrm{s}}^{\mathrm{A}}$-th pixel in the coordinate system of BS A to the $n_{\mathrm{s}}^{\mathrm{B}}$-th pixel in the BS B coordinate system is expressed as $h_{n_{\mathrm{f}}}^{\mathrm{UE}n_{\mathrm{T}} \to \mathrm{P}n_{\mathrm{s}}^{\mathrm{B}} \to \mathrm{BS}n_{\mathrm{R}}} / h_{n_{\mathrm{f}}}^{\mathrm{UE}n_{\mathrm{T}} \to \mathrm{P}n_{\mathrm{s}}^{\mathrm{A}} \to \mathrm{BS}n_{\mathrm{R}}}$ based on (2) where the global coordinate system is replaced by the coordinate system from the view of BSs A and B respectively.

Considering the above two mapping coefficients, for the n_{f}-th subcarrier and the n_{R}-th UE, we map the BS A coordinate system $\boldsymbol{x}^{\mathrm{BSA}}$ to the BS B coordinate system $\boldsymbol{x}^{\mathrm{BSB}}$, i.e.,

$$\boldsymbol{x}^{\mathrm{BSB}} = \mathbf{K}_{n_{\mathrm{f}},n_{\mathrm{R}}}^{\mathrm{BSA}\to\mathrm{BSB}} \boldsymbol{x}^{\mathrm{BSA}}, \quad (8)$$

where $\mathbf{K}_{n_{\mathrm{f}},n_{\mathrm{R}}}^{\mathrm{BSA}\to\mathrm{BSB}} \in \mathbb{C}^{N_{\mathrm{s}}^{\mathrm{B}} \times N_{\mathrm{s}}^{\mathrm{A}}}$ is the mapping matrix. This matrix maps the pixel-wise propagation gains from the view of BS A to that of BS B, accounting for both the power projection and the phase alignment between the two views. Each element is defined as:

$$k_{n_{\mathrm{f}},n_{\mathrm{R}}}^{\mathrm{BSA}\to\mathrm{BSB}}\left(n_{\mathrm{s}}^{\mathrm{B}}, n_{\mathrm{s}}^{\mathrm{A}}\right) = \frac{S_{n_{\mathrm{s}}^{\mathrm{A}},n_{\mathrm{s}}^{\mathrm{B}}}}{S_{n_{\mathrm{s}}^{\mathrm{A}}}} \frac{h_{n_{\mathrm{f}}}^{\mathrm{UE}n_{\mathrm{T}} \to \mathrm{P}n_{\mathrm{s}}^{\mathrm{B}} \to \mathrm{BS}n_{\mathrm{R}}}}{h_{n_{\mathrm{f}}}^{\mathrm{UE}n_{\mathrm{T}} \to \mathrm{P}n_{\mathrm{s}}^{\mathrm{A}} \to \mathrm{BS}n_{\mathrm{R}}}}, \quad (9)$$

and $N_{\mathrm{s}}^{\mathrm{A}}$, $N_{\mathrm{s}}^{\mathrm{B}}$ are the number of pixels from BS A's view and the number of pixels from BS B's view respectively. In (9), the first term represents the power mapping coefficient and the second term represents the phase mapping coefficient between the two corresponding propagation paths. This mapping matrix shows how spatial energy contributions and phase information are transferred between different BS views. It can be concluded that the sum of each column of the mapping matrix $\mathbf{K}_{n_{\mathrm{f}},n_{\mathrm{R}}}^{\mathrm{BSA}\to\mathrm{BSB}}$ is 1, because the sum of the overlapping areas of a pixel and other pixels is equal to the area of the pixel itself. The power mapping coefficient from $n_{\mathrm{s}}^{\mathrm{A}}$-th pixel to $n_{\mathrm{s}}^{\mathrm{B}}$-th pixel described in (9) is denoted by $S_{n_{\mathrm{s}}^{\mathrm{A}},n_{\mathrm{s}}^{\mathrm{B}}} / S_{n_{\mathrm{s}}^{\mathrm{A}}}$.

In addition, the coordinate system mapping method expressed in (8) is also suitable for mapping UE to UE and UE to BS. The calculation of the mapping matrix is also related to the coordinate values that need to be mapped. For example, mapping the scattering coefficient of a pixel only requires a power mapping coefficient, while mapping the propagation gain of a pixel also requires a phase mapping coefficient.

C. Multi-View ISAC Model

Multiple UEs and BSs in the environment share time-frequency resources, we assume that the channel is a quasi-static channel where transmission time T_{b} of a communication resource block is much shorter than the channel coherence time T_{c}, i.e. $T_{\mathrm{b}} \ll T_{\mathrm{c}}$, and the UE is stationary within the coherence time.

Our research focuses on preserving the communication process and implementing sensing passively, using only existing communication signals and equipment to avoid interruptions in communication services. In the uplink, we adopt a standard OFDM-based multi-subcarrier signaling scheme, where each UE transmits modulated data symbols over orthogonal subcarriers with fixed subcarrier spacing and a cyclic prefix. Under this communication model, practical issues such as frequency-selective fading and inter-carrier interference are already handled by the communication system through channel estimation, equalization, and OFDM orthogonality [29]. The estimated channel state information (CSI) on all subcarriers is then passively exploited for sensing, enabling joint multi-view sensing without introducing additional signaling overhead. For the n_{f}-th subcarrier, the estimated multipath propagation gain $\hat{h}_{n_{\mathrm{f}}}^{\mathrm{UE1}\to\mathrm{BS1}}$ from UE 1 to BS 1 is expressed as

$$\begin{aligned} &\hat{h}_{n_{\mathrm{f}}}^{\mathrm{UE1}\to\mathrm{BS1}} \\ &= \left(\tilde{\boldsymbol{h}}_{n_{\mathrm{f}}}^{\mathrm{UE1}\to\mathrm{P}N_{\mathrm{s}}^{\mathrm{UE1}}\to\mathrm{BS1}} \odot \boldsymbol{v}^{\mathrm{UE1}}\right) \mathbf{K}_{n_{\mathrm{f}}}^{\mathrm{BS1}\to\mathrm{UE1}} \boldsymbol{x}^{\mathrm{BS1}} + n, \quad (10) \\ &= \left(\tilde{\boldsymbol{h}}_{n_{\mathrm{f}}}^{\mathrm{UE1}\to\mathrm{P}N_{\mathrm{s}}^{\mathrm{BS1}}\to\mathrm{BS1}} \odot \boldsymbol{v}^{\mathrm{BS1}}\right) \mathbf{K}_{n_{\mathrm{f}}}^{\mathrm{BS1}\to\mathrm{BS1}} \boldsymbol{x}^{\mathrm{BS1}} + n, \quad (11) \end{aligned}$$

$$\begin{bmatrix} \hat{h}_{n_\mathrm{f}}^{\mathrm{UE1}\to\mathrm{BS1}} \\ \vdots \\ \hat{h}_{n_\mathrm{f}}^{\mathrm{UE}N_\mathrm{T}\to\mathrm{BS}N_\mathrm{R}} \end{bmatrix}_{N_\mathrm{R}N_\mathrm{T}\times 1} = \begin{bmatrix} \big(\tilde{\boldsymbol{h}}_{n_\mathrm{f}}^{\mathrm{UE1}\to\mathrm{P}N_\mathrm{s}^{\mathrm{UE1}}\to\mathrm{BS1}} \odot \boldsymbol{v}^{\mathrm{UE1}}\big)\mathbf{K}_{n_\mathrm{f}}^{\mathrm{BS1}\to\mathrm{UE1}} \\ \vdots \\ \big(\tilde{\boldsymbol{h}}_{n_\mathrm{f}}^{\mathrm{UE}N_\mathrm{T}\to\mathrm{P}N_\mathrm{s}^{\mathrm{UE}N_\mathrm{T}}\to\mathrm{BS}N_\mathrm{R}} \odot \boldsymbol{v}^{\mathrm{UE}N_\mathrm{T}}\big)\mathbf{K}_{n_\mathrm{f}}^{\mathrm{BS1}\to\mathrm{UE}N_\mathrm{T}} \end{bmatrix}_{N_\mathrm{R}N_\mathrm{T}\times N_\mathrm{s}^{\mathrm{BS1}}} \boldsymbol{x}^{\mathrm{BS1}} + \boldsymbol{n}. \quad (13)$$

where the results of (10) and (11) are the same, but they are the observation equations[8] from the BS1 and UE1 views, respectively and n represents the noise[9]. The mapping matrix from UE 1 to BS 1 is represented by $\mathbf{K}^{\mathrm{BS1}\to\mathrm{UE1}} \in \mathbb{C}^{N_\mathrm{s}^{\mathrm{UE1}}\times N_\mathrm{s}^{\mathrm{BS1}}}$. In (10), We map the coordinate system of the UE1 view to the coordinate system of the BS1 view where $\tilde{\boldsymbol{h}}_{n_\mathrm{f}}^{\mathrm{UE1}\to\mathrm{P}N_\mathrm{s}^{\mathrm{UE1}}\to\mathrm{BS1}} \in \mathbb{C}^{1\times N_\mathrm{s}^{\mathrm{UE1}}}$ is the free space propagation gain from the UE 1 to the BS 1 which scattered by $N_\mathrm{s}^{\mathrm{UE1}}$ pixels. The occlusion vector[10] from UE 1 to the BS 1 is denoted as $\boldsymbol{v}^{\mathrm{UE1}} \in \{0,1\}^{1\times N_\mathrm{s}^{\mathrm{UE1}}}$,

$$\boldsymbol{v}^{\mathrm{UE1}} = f_\mathrm{o}(\boldsymbol{x}, \boldsymbol{p}^{\mathrm{UE1}}), \quad (12)$$

where $\boldsymbol{p}^{\mathrm{UE1}}$ is the position of pixels from UE1's view. Environmental information $\boldsymbol{x}$ includes $\boldsymbol{x}^{\mathrm{BS1}}$, $\boldsymbol{x}^{\mathrm{UE1}}$, etc. $f_\mathrm{o}(\cdot)$ which is defined in (3)-(5) in [8] based on the geometric relationship. However, in this paper, the occlusion detection function can be simplified to the intuitive occlusion detection method described in section II-B, that is, detecting objects in closer pixels that occlude farther pixels. (12) shows that the occlusion vector is determined by the environmental target information and the divided pixel position. The distribution of different occlusion vectors reflects different partial views from different UEs and BSs. The zero element in the vector means that the UE or the BS cannot sense the corresponding pixel, and the corresponding pixel is out of the sensing range of this partial view.

In (11), similar to (10), $\tilde{\boldsymbol{h}}_{n_\mathrm{f}}^{\mathrm{UE1}\to\mathrm{PN}_\mathrm{s}^{\mathrm{BS1}}\to\mathrm{BS1}} \in \mathbb{C}^{1\times N_\mathrm{s}^{\mathrm{BS1}}}$ is the free space propagation gain from the UE 1 to the BS 1 which scattered by $N_\mathrm{s}^{\mathrm{BS1}}$ pixels. The occlusion vector from UE 1 to the BS 1 is represented by $\boldsymbol{v}^{\mathrm{BS1}} \in \{0,1\}^{1\times N_\mathrm{s}^{\mathrm{BS1}}}$ and $\mathbf{K}^{\mathrm{BS1}\to\mathrm{BS1}} \in \mathbb{C}^{N_\mathrm{s}^{\mathrm{BS1}}\times N_\mathrm{s}^{\mathrm{BS1}}}$ is the mapping matrix from BS 1 to BS 1. There is a theoretical mapping method for $\boldsymbol{v}^{\mathrm{UE1}}$ and $\boldsymbol{v}^{\mathrm{BS1}}$, $\boldsymbol{v}^{\mathrm{UE1}} = \mathbb{1}(\mathbf{K}^{\mathrm{BS1}\to\mathrm{UE1}}\boldsymbol{v}^{\mathrm{BS1}} \neq 0)$, where $\mathbb{1}(\cdot)$ means the corresponding element value is 1. However, the occlusion effect is determined by the environment, and the occlusion vector $\boldsymbol{v}^{\mathrm{UE1}}$ or $\boldsymbol{v}^{\mathrm{BS1}}$ represents the comprehensive result of occlusion from the two views of BS and UE, which should be jointly detected through the imaging results from different views instead of simply mapping the occlusion vector of one view to the occlusion vector of another view.

In this paper, we take the BS1 view coordinate system as an example of how to design the sensing algorithm. The proposed algorithm in the practical scenario is suitable for any UE and BS. Based on (10), we jointly process the multipath propagation gain of N_T UEs and N_R BSs, which results in (13) as shown at the top of this page. By stacking the estimated channels over all subcarriers, transmitters, and receivers, the multi-subcarrier sensing model can be written as (14):

$$\hat{\mathbf{H}} = \mathbf{A}^{\mathrm{BS1}}\boldsymbol{x}^{\mathrm{BS1}} + \boldsymbol{n}. \quad (14)$$

where $\hat{\mathbf{H}} = \big[\hat{\boldsymbol{h}}_1; \ldots; \hat{\boldsymbol{h}}_{N_\mathrm{f}}\big]$, $\mathbf{A}^{\mathrm{BS1}} = \big[\mathbf{A}_1^{\mathrm{BS1}}; \ldots; \mathbf{A}_{N_\mathrm{f}}^{\mathrm{BS1}}\big]$. $\hat{\boldsymbol{h}}_{n_\mathrm{f}}$ is the channel estimation result obtained in the communication which is the left term of (13), and $\mathbf{A}_{N_\mathrm{f}}^{\mathrm{BS1}}$ is the reference matrix in the right term of (13). $\boldsymbol{x}^{\mathrm{BS1}}$ represents the unknown environment information to be solved and $\boldsymbol{v}^{\mathrm{UE1}}, \ldots, \boldsymbol{v}^{\mathrm{UE}N_\mathrm{T}}$ are the unknown occlusion vectors. The remaining variables can be obtained through the models in III-A and III-B.

IV. MULTI-VIEW ENVIRONMENT SENSING ALGORITHM

In this section, we propose an iterative multi-view environment sensing algorithm. Fig. 5 illustrates the proposed multi-view ISAC scenario where each BS first acquires its local CSI from the uplink transmissions of UEs. To perform multi-view sensing, each BS also requires the CSI collected at the other BSs.[11] Once this information is exchanged, each BS can independently execute the proposed iterative multi-view sensing algorithm using its own pixel grid and mapping matrix. In this way, the full multi-view sensing can be achieved locally at each BS without relying on a single central processing node. we jointly consider multi-view measurements of the environment and transform the multi-view sensing problem into a CS sparse reconstruction problem and propose a factor graph relationship between multi-view measurements and environmental information based on an approximate message passing algorithm, and the occlusion effect is solved by the AO algorithm. It is worth noting that in the following derivations, for clarity, we omit all superscripts related to "BS1" unless otherwise specified. For example, (14) would be written as $\hat{\mathbf{H}} = \mathbf{A}\boldsymbol{x} + \boldsymbol{n}$.

A. Problem Formulation

As shown in Fig. 5, we let the uplink communication process between a transceiver pair (n_T-th UE and n_R-th BS) be a measurement of one view and obtain channel estimation result, which is modeled in (13). Jointly processing the

---

[8] Note that for clarity, we do not introduce multiple-input and multiple-output (MIMO) technology into the model. The transceiver in this paper can be extended to a multi-antenna system, and the proposed method is still effective. For specific models that include MIMO and beam-space modeling, please refer to previous work [30].

[9] The noise comes from the noise of the channel itself, and also includes the channel estimation error caused by clock bias, phase bias and UE positioning errors. In practical systems, calibration methods should be used to reduce these errors.

[10] The binary occlusion detection vector may be affected by partial occlusion, which typically occurs only near object boundaries and is uncommon in most regions. As shown in Section VI, the proposed algorithm remains effective even under such conditions.

[11] This exchange of CSI is the additional data exchange for multi-view sensing, and its communication overhead is comparable to that of centralized processing, where a central node collects CSI from all BSs.

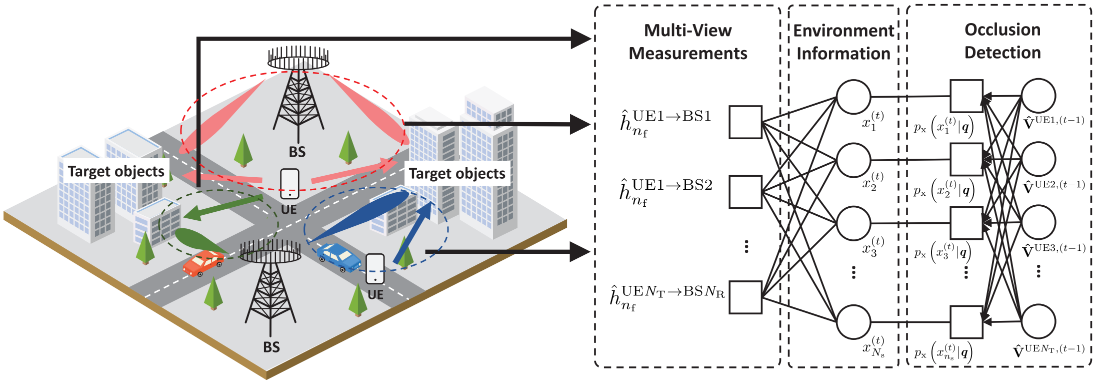


Fig. 5. The proposed multi-view ISAC scenario and the factor graph.

propagation gains of N_f subcarriers between multiple UEs and BSs, we can obtain a total of M measurements, where we let $M = N_\mathrm{f}N_\mathrm{T}N_\mathrm{R}$ to simplify the expression. However, even if the resources of multiple transceivers and multiple subcarriers are exploited, environmental information still has more unknown variables than the number of measurements. Therefore, based on the sparsity of environmental target information, we transform the environment sensing equation (14) into a CS reconstruction optimization problem:

$$\min_{\boldsymbol{x}} \quad \|\boldsymbol{x}\|_1, \tag{15a}$$

$$\text{s.t.} \quad \|\hat{\mathbf{H}} - \mathbf{A}\boldsymbol{x}\|_2 \leq \varepsilon, \tag{15b}$$

where $\boldsymbol{x}$ is independent of frequency. We do not focus on optimizing the occlusion vector here, as our primary objective is to sense the environment. Additionally, the occlusion vector remains independent of frequency, ε is a slack variable.

B. Proposed Alternative Optimization Algorithm

In this paper, the GAMP algorithm [23] is used to solve the CS reconstruction problem in (15) by using an iterative decomposition method as shown in Algorithm 1 in [4]. To this end, we formulate a prior density of the scattering coefficient. According to the sparsity of the environment, we model $\boldsymbol{x}$ to follow a Bernoulli-Gaussian distribution in a limited interval, defined as

$$p_\mathrm{x}\left(x_{n_\mathrm{s}}|\boldsymbol{q}\right) = (1-\alpha+\beta)\delta\left(x_{n_\mathrm{s}}\right) + \alpha\mathcal{N}\left(x_{n_\mathrm{s}}|\theta^\mathrm{x},\sigma^\mathrm{x}\right)\big(u(x_{n_\mathrm{s}}) - u(x_{n_\mathrm{s}}-1)\big), \tag{16}$$

where $\boldsymbol{q} \triangleq [\alpha,\beta,\theta^\mathrm{x},\sigma^\mathrm{x}]$ denotes all parameters of the Bernoulli-Gaussian distribution, $\delta(\cdot)$ is the Dirac function, $u(\cdot)$ is the step function, α is the sparsity coefficient of the environment, $\beta{=}\int_{x\in(-\infty,0]\cup[1,+\infty)}\alpha\mathcal{N}\left(x|\theta^\mathrm{x},\sigma^\mathrm{x}\right)\mathrm{d}x$, and $\theta^\mathrm{x} \in [0,1]$ and σ^x represent the mean and the standard deviation of the environmental information $\boldsymbol{x}$, respectively.

Different from traditional CS reconstruction algorithms, in the considered multi-view sensing scenario, the CS measurement matrix is an unknown matrix $\mathbf{A}$. The key point of the algorithm design is that we need to solve both the unknown environment target variables and the occlusion relationship. For optimization problem (15), the traditional GAMP algorithm ignores the occlusion matrix $\mathbf{V} \triangleq \{\boldsymbol{v}^{\mathrm{UE1}},\ldots,\boldsymbol{v}^{\mathrm{UE}N_\mathrm{T}}\}$ and regards the CS measurement matrix $\mathbf{A}$ as a known matrix, which makes the result inaccurate. Compared with GAMP, the bilinear GAMP algorithm can solve the bilinear CS problem. It solves the unknown CS measurement matrix $\mathbf{A}$ and the sparse variable $\boldsymbol{x}$ according to the linear relationship between the two variables. However, the bilinear GAMP algorithm does not utilize the relationship between the occlusion matrix $\mathbf{V}$ and the environmental information $\boldsymbol{x}$ and lacks sufficient prior information for the algorithm to converge to an optimal solution. Considering the occlusion effect in a multi-view sensing scenario, the GAMP-MVSVR algorithm proposed in [8] jointly solves the occlusion matrix $\mathbf{V}$ and the environmental information $\boldsymbol{x}$. However, the centralized algorithm is not suitable for the proposed distributed scenario, and it is not easy to accurately solve the factor graph relationship for the proposed complex system model. In this section, we consider the AO strategy for algorithm design to solve the shortcomings of the above algorithms.

Theoretically, in (15), the occlusion matrix $\mathbf{V}$ contains unknown discrete variables. By traversing the entire feasible domain, the optimal occlusion matrix $\mathbf{V}$ can be found, the environmental information $\boldsymbol{x}$ can be solved, and the upper bound of the sensing performance can be achieved. Undoubtedly, the complexity of the traversal algorithm is impractical. In the proposed message-passing-based sensing algorithm, the main challenge in achieving the theoretical global optimum is that the measurement matrix $\mathbf{A}$ in (15) includes the unknown occlusion matrix $\mathbf{V}$, and that $\mathbf{V}$ is highly correlated with the environmental information $\boldsymbol{x}$. This coupling makes the proposed problem substantially more challenging than the conventional CS problem.

To overcome the strong coupling between the environmental information $\boldsymbol{x}$ and the occlusion matrix $\mathbf{V}$, we reformulate (15) as a joint multi-variable optimization problem, where both the environmental information $\boldsymbol{x}$ and the occlusion matrix $\mathbf{V}$ are treated as optimization variables, as expressed in (17). This formulation introduces a trade-off between the sparsity of $\boldsymbol{x}$ and the consistency of the occlusion detection function $f_\mathrm{o}(\cdot)$,

$$\min_{\boldsymbol{x},\boldsymbol{v}^{\mathrm{UE}n_\mathrm{T}}} \|\boldsymbol{x}\|_1 + (1-\eta)\|\mathbf{V}^{\mathrm{UE}n_\mathrm{T}} - f_\mathrm{o}(\boldsymbol{x},\boldsymbol{p}^{\mathrm{UE}n_\mathrm{T}})\|_\mathrm{F}^2 \tag{17a}$$

$$\text{s.t.} \quad \|\hat{\mathbf{H}} - \mathbf{A}\boldsymbol{x}\|_2 \leq \varepsilon, \tag{17b}$$

where

$$\mathbf{A}_{n_\mathrm{f}}(n_\mathrm{T}, n_\mathrm{R}) = \left(\tilde{\boldsymbol{h}}_{n_\mathrm{f}}^{\mathrm{UE}n_\mathrm{T}\to\mathrm{P}N_\mathrm{s}^{\mathrm{UE}n_\mathrm{T}}\to\mathrm{BS}n_\mathrm{R}} \odot \boldsymbol{v}^{\mathrm{UE}n_\mathrm{T}}\right)\mathbf{K}_{n_\mathrm{f}}^{\mathrm{BS1}\to\mathrm{UE}n_\mathrm{T}}. \tag{18}$$

In addition, η is a given weight, $\|\cdot\|_\mathrm{F}$ represents the Frobenius norm, $n_\mathrm{T} = 1, \ldots, N_\mathrm{T}$, and $n_\mathrm{R} = 1, \ldots, N_\mathrm{R}$. For the optimization problem in (17), considering the relationship between $\boldsymbol{x}$ and $\mathbf{V}$, we employ an alternating optimization (AO) strategy, in which $\boldsymbol{x}$ and $\mathbf{V}$ are updated iteratively to to determine an appropriate value of η and obtain an efficient solution. In the t-th iteration, we decompose the optimization problem in (17) into the following two optimization sub-problems.

- Sparse reconstruction sub-problem: Given the measurement vector based on the occlusion detection results $\hat{\boldsymbol{v}}^{\mathrm{UE}n_\mathrm{T},(t-1)}$ in the $(t-1)$-th iteration, $\hat{\boldsymbol{x}}^{(t)}$ is reconstructed as the solution of a constrained ℓ_1-minimization problem,

  $$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1, \tag{19a}$$

  $$\text{s.t.} \;\; \|\hat{\mathbf{H}} - \hat{\mathbf{A}}^{(t-1)}\boldsymbol{x}\|_2 \leq \varepsilon, \tag{19b}$$

  where

  $$\hat{\mathbf{A}}_{n_\mathrm{f}}^{(t-1)}(n_\mathrm{T}, n_\mathrm{R}) = \left(\tilde{\boldsymbol{h}}_{n_\mathrm{f}}^{\mathrm{UE}n_\mathrm{T}\to\mathrm{P}N_\mathrm{s}^{\mathrm{UE}n_\mathrm{T}}\to\mathrm{BS}n_\mathrm{R}} \odot \hat{\boldsymbol{v}}^{\mathrm{UE}n_\mathrm{T},(t-1)}\right)\mathbf{K}_{n_\mathrm{f}}^{\mathrm{BS1}\to\mathrm{UE}n_\mathrm{T}}. \tag{20}$$

  The GAMP algorithm is used to perform the t-th CS sparse reconstruction. Based on (16), To further reduce the dimensionality of the problem and enhance sparsity, we exploit the occlusion matrix to set a structured prior on $\boldsymbol{x}$ according to the $(t-1)$-th occlusion detection result,

  $$p_\mathrm{x}\left(x_{n_\mathrm{s}}^{(t)}|\boldsymbol{q}\right) = \begin{cases} 0, \;\; \text{if} \;\; \sum_{N_\mathrm{T}} \hat{\boldsymbol{v}}^{\mathrm{UE}n_\mathrm{T},(t-1)} \leq \gamma^v, \\ p_\mathrm{x}\left(x_{n_\mathrm{s}}|\boldsymbol{q}\right), \;\; \text{Otherwise}, \end{cases} \tag{21}$$

  where $p_\mathrm{x}\left(x_{n_\mathrm{s}}|\boldsymbol{q}\right)$ is defined in (16) and γ^v is a threshold. The threshold γ^v should be set to a small value close to 0. In (21), we detect whether the n_s-th pixel is completely occluded from the view of all transceivers according to the occlusion matrix. This treatment suppresses irrelevant variables and improves the accuracy and stability of reconstruction. We initialize the occlusion vector $\boldsymbol{v}$ as an all-one vector to avoid missing any potential scatterers during the optimization process.
- Occlusion detection sub-problem: From a general perspective, the occlusion detection sub-problem aims to estimate the occlusion matrix $\hat{\boldsymbol{v}}^{\mathrm{UE}n_\mathrm{T},(t)}$ based on the current environmental estimate $\hat{\boldsymbol{x}}^{(t)}$ while satisfying the channel propagation gain constraint, which can be expressed as

  $$\min_{\boldsymbol{v}^{\mathrm{UE}n_\mathrm{T}}} \|\boldsymbol{v}^{\mathrm{UE}n_\mathrm{T}} - f_\mathrm{o}(\hat{\boldsymbol{x}}^{(t)}, \boldsymbol{p}^{\mathrm{UE}n_\mathrm{T}})\|_\mathrm{F}^2 \tag{22a}$$

  $$\text{s.t.} \;\; \|\tilde{\mathbf{H}} - \mathbf{A}\hat{\boldsymbol{x}}^{(t)}\|_2 \leq \varepsilon. \tag{22b}$$

  Since the occlusion matrix $\mathbf{V}$ is uniquely determined by the spatial distribution of the environment $\boldsymbol{x}$, the solution to (22) depends primarily on the current estimate $\hat{\boldsymbol{x}}^{(t)}$. Existing works typically rely on exhaustive search, traversing all possible pixel combinations that might occlude one another [8]. However, this approach suffers from high computational complexity. In our work, benefiting from the pixel-based discretization method introduced in Section II-B, we are able to design a computationally efficient method that greatly reduces the complexity compared with a centralized exhaustive search. Specifically, based on II-B, we use (23) and (24) to directly update the occlusion matrix via a thresholded version of the estimated environmental information,

  $$\hat{\boldsymbol{v}}^{\mathrm{UE}n_\mathrm{T},(t)} = f_\mathrm{o}(\tilde{\boldsymbol{x}}^{(t)}, \boldsymbol{p}^{\mathrm{UE}n_\mathrm{T}}), \tag{23}$$

  $$\tilde{x}_{n_\mathrm{s}}^{(t)} = \{\hat{x}_{n_\mathrm{s}}^{(t)} | \hat{x}_{n_\mathrm{s}}^{(t)} \geq \gamma^x\}, \tag{24}$$

  where $n_\mathrm{s} = 1, \ldots, N_\mathrm{s}$ and $\tilde{\boldsymbol{x}}^{(t)}$ is obtained by applying an element-wise indicator function to $\hat{\boldsymbol{x}}^{(t)}$. Each pixel is either kept or suppressed depending on whether its estimated scattering coefficient exceeds a threshold $\gamma^x \in [0, 1]$. This step serves to filter out noise and weak artifacts before occlusion detection, thereby improving robustness. Due to the complexity of the proposed model, deriving a theoretical error bound or convergence guarantee based on message-passing theory is theoretically difficult. Therefore, the thresholds are selected empirically. To improve convergence, the threshold γ^x is gradually decreased across iterations, ensuring that constraint (22b) is satisfied and that the algorithm converges to a physically consistent solution. In complex scenarios, it is recommended to use a relatively high initial value for γ^x to avoid overly aggressive detection, which could deteriorate convergence.

In the proposed AO algorithm, we perform optimization problems (19) and (22) alternately, and the results of the two optimization problems promote each other and converge to accurate sensing results. The specific AO algorithm is summarized in Algorithm 1.

C. Convergence and Computational Complexity Analysis

1) Convergence Analysis: The proposed AO algorithm consists of two sub-optimization problems as (19) and (22). For optimization problem (19), assume that the measurement matrix $\mathbf{A}_{M\times N}$ is an independent identically distributed Gaussian random matrix with infinite dimension and fixed aspect ratio, i.e., $M, N \to \infty, M/N = \zeta$. In this case, it is proved that the asymptotic behavior of GAMP obeys the state evolution equation, and the iterative estimation is consistent with the true value. Therefore, in (19), when the prior information (21) is accurate, a sufficient number of BSs and UEs at random locations can ensure the convergence of the algorithm.

For optimization problem (22), when the environmental information $\boldsymbol{x}$ is completely accurate, the optimal solution of (22) is to directly calculate the occlusion as (23) and (24) according to the given $\boldsymbol{x}$. However, the environmental information $\boldsymbol{x}$ is only an estimated value when the proposed AO algorithm is executed. Therefore, (19) only obtains a local

Algorithm 1 The Proposed AO Algorithm

```
Input: Number of UEs and BSs required for the sensing method, i.e., $N_\mathrm{T}$ and $N_\mathrm{R}$. The pixel division parameters $\theta_\mathrm{s}$ and $d_\mathrm{s}$ of each BS and UE. Number of carriers $N_\mathrm{f}$ and their wavelengths $\lambda_{N_\mathrm{T}}$. The channel propagation gain $\tilde{\mathbf{H}}$ estimated by BSs in (17).
Initialization: Pre-dividing pixel coordinate systems for BSs and UEs which denoted as $\boldsymbol{x}^{\mathrm{BS1}},\ldots,\boldsymbol{x}^{\mathrm{BS}N_\mathrm{R}}$ and $\boldsymbol{x}^{\mathrm{UE1}},\ldots,\boldsymbol{x}^{\mathrm{UE}N_\mathrm{T}}$. Calculate $\mathbf{K}_{n_\mathrm{f}}^{\mathrm{UE1}\to\mathrm{BS1}},\ldots,\mathbf{K}_{n_\mathrm{f}}^{\mathrm{UE}N_\mathrm{T}\to\mathrm{BS1}}$ in (17) from (9). Calculate $\tilde{\mathbf{H}}_{n_\mathrm{f},N_\mathrm{s}^{\mathrm{UE}N_\mathrm{T}}}^{\mathrm{UE}N_\mathrm{T}\to\mathrm{BS}N_\mathrm{R}}$ in (17) from (2). Set $\boldsymbol{v}^{\mathrm{UE}n_\mathrm{T},(0)} = \mathbf{1}$.
while $\|\boldsymbol{x}^{\mathrm{BS1},(t)} - \boldsymbol{x}^{\mathrm{BS1},(t-1)}\|_2^2 > \varepsilon_t \|\boldsymbol{x}^{\mathrm{BS1},(t-1)}\|_2^2$, where $\varepsilon_t$ is a given error tolerance value in $t$-th iteration. do
    Calculate the prior probability of $\boldsymbol{x}^{\mathrm{BS1},(t)}$ in (21).
    Solve optimization problem (19) by GAMP algorithm.
    Solve optimization problem (22) by (23) and (24).
    $t = t + 1$.
end while
Output: Estimated target object $\boldsymbol{x}^{\mathrm{BS1},(t)}$ in BS1 view coordinate system.
```

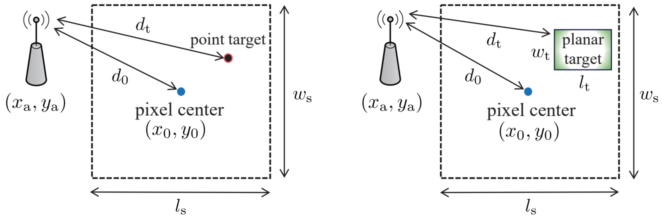


Fig. 6. Two cases of pixel division errors caused by incomplete target filling. (left: point target case; right: planar target case).

optimal solution, and the result may deviate from the global optimal solution, which depends on the difference between the estimated value $\boldsymbol{x}$ and the actual result.

Combining the above two aspects of analysis, from the perspective of alternating iteration, the solutions to the two sub-problems are both locally optimal, and the algorithm cannot certainly converge to the global optimal solution. In addition, to ensure flexibility, scalability and practical performance, we improve the convergence of the algorithm in two ways. First, through multi-view observations, we obtain as many measurement equations as possible, thereby increasing the constraints on the environmental information $\boldsymbol{x}$. Second, by accurately modeling the geometric occlusion relationships among pixels, we constrain the feasible solution space of $\mathbf{V}$ and $\boldsymbol{x}$, which promotes consistent and stable reconstruction. It is currently difficult to provide a rigorous theoretical proof that the proposed algorithm converges to the global optimum. The proposed approach is evaluated based on consistent performance gains over existing methods in extensive numerical experiments. These results demonstrate that, despite the lack of strong theoretical guarantees, the proposed algorithm is reliable and effective in practical scenarios.

2) Computational Complexity Analysis: The computational complexity of the two sub-problems is analyzed as follows: The computational complexity of the CS reconstruction in (19) is expressed as $\mathcal{O}\left(N_\mathrm{s}N_\mathrm{f}N_\mathrm{T}N_\mathrm{R}\right)$. It is considered that the number of pixels N_s has a major influence on the computational complexity. The proposed method can use pixels with large areas, thereby reducing the number of pixels and reducing the resources required for sensing. For optimization problem (22), benefiting from the pre-division of the coordinate system during initialization, the occlusion detection problem can be quickly calculated by table lookup. Compared with the computational complexity $\mathcal{O}\left(N_\mathrm{T}N_\mathrm{R}N_\mathrm{s}\|\boldsymbol{x}\|_0\right)$ of the centralized approach, the proposed occlusion detection method has a complexity of only $\mathcal{O}\left(\|\boldsymbol{x}\|_0\right)$, making the method scalable and more suitable for large-scale deployment. In summary, the complexity of the proposed AO algorithm is expressed as $\mathcal{O}\left(N_\mathrm{s}N_\mathrm{f}N_\mathrm{T}N_\mathrm{R} + \|\boldsymbol{x}\|_0\right)$. For the pixel coordinate position calculation and the numerical integration of the propagation gain in the initialization stage, not only can they be executed in parallel, but they only need to be calculated once in advance before the algorithm is executed, so the impact on the computational complexity is small.

3) Communication Overhead Analysis: In the distributed multi-view sensing framework, each BS first estimates its local CSI from uplink transmissions and then exchanges it with other cooperating BSs to enable multi-view reconstruction. This CSI exchange constitutes the main source of communication overhead. The overall communication overhead for sensing is $\mathcal{O}(N_\mathrm{T}N_\mathrm{R}N_\mathrm{f})$. Unlike imaging methods that require sharing raw measurements or high-dimensional sensing results, the proposed framework only exchanges CSI, whose dimension does not grow with the sensing resolution. As a result, increasing the sensing area or refining the pixel grid does not lead to additional communication burden, which enables scalable operation in large-scale scenarios. Moreover, the proposed architecture allows flexible cooperation among BSs and UEs according to practical network conditions. CSI exchange can be restricted to neighboring nodes or selected clusters, thereby avoiding full network-wide information sharing. This further limits communication overhead and reduces sensitivity to network latency.

V. SYSTEM PERFORMANCE ANALYSIS

A. Pixel Division Accuracy

In this section, we analyze the impact of pixel division on sensing error and verify that the proposed pixel model represents the real environment more accurately. In this section we assume that pixels are rectangular to make the formula clear. As shown in Fig. 6, we analyze the errors caused by pixel division in two cases: *point targets* and *planar targets*. Let the antenna be located at $(x_\mathrm{a}, y_\mathrm{a})$ and the central of the pixel at (x_0, y_0) and introduce $d(x, y) = \sqrt{(x_\mathrm{a} - x)^2 + (y_\mathrm{a} - y)^2}$ as the distance between the antenna and an arbitrary point in the pixel. We will use $d(x, y)$, $d(u, v)$ or $d(\tilde{u}, \tilde{v})$ with identical meaning but different dummy variables.

1) Point Targets: For the point target shown in 6(a), modeled using conventional computational imaging as in (1), the phase error $e_{1,\mathrm{c}}$ caused by the point target deviating from the pixel center, is derived by calculating the average error

$$e_{1,\mathrm{c}} = \int_{x_0-l_\mathrm{s}/2}^{x_0+l_\mathrm{s}/2} \int_{y_0-w_\mathrm{s}/2}^{y_0+w_\mathrm{s}/2} \frac{2\pi |d(x,y) - d_0|}{l_\mathrm{s} w_\mathrm{s} \lambda} \mathrm{d}x \mathrm{d}y, \tag{25}$$

where $d_0 = \sqrt{(x_\mathrm{a} - x_0)^2 + (y_\mathrm{a} - y_0)^2}$ is the distance between the antenna and the pixel.

The phase error $e_{1,\mathrm{p}}$ caused by the proposed method is expressed as

$$e_{1,\mathrm{p}} = \int_{x_0-l_\mathrm{s}/2}^{x_0+l_\mathrm{s}/2} \int_{y_0-w_\mathrm{s}/2}^{y_0+w_\mathrm{s}/2} \frac{2\pi |d(x,y) - d_\mathrm{p}|}{l_\mathrm{s} w_\mathrm{s} \lambda} \mathrm{d}x \mathrm{d}y, \tag{26}$$

where d_p is the average distance from pixel to antenna in an integrated form

$$d_\mathrm{p} = \int_{x_0-l_\mathrm{s}/2}^{x_0+l_\mathrm{s}/2} \int_{y_0-w_\mathrm{s}/2}^{y_0+w_\mathrm{s}/2} \frac{d(u,v)}{l_\mathrm{s} w_\mathrm{s}} \mathrm{d}u \mathrm{d}v. \tag{27}$$

Compared to the conventional computational imaging method, the proposed method has similar errors for point targets. This is because, for point targets, the average distance error caused by random distribution within pixels is similar to the average distance error obtained by the proposed integration method. For square pixels, the average distance d_p calculated in (27) is almost the same as the distance d_0 from the center of the pixel to the antenna in (25). However, ideal point targets are unusual in practical scenarios. In a practical scenario, smaller targets can be accurately sensed by dividing smaller pixels until the resolution limit is reached, which is determined by the wavelength.

2) Planar Targets: For the planar target as shown in Fig. 6(b) , the phase error $e_{2,\mathrm{c}}$ caused by the conventional method is expressed as

$$e_{2,\mathrm{c}} = \int_{x_0-l_\mathrm{s}/2}^{x_0+l_\mathrm{s}/2} \int_{y_0-w_\mathrm{s}/2}^{y_0+w_\mathrm{s}/2} \frac{2\pi |d_\mathrm{t}(x,y) - d_0|}{l_\mathrm{s} w_\mathrm{s} \lambda} \mathrm{d}x \mathrm{d}y, \tag{28}$$

$$d_\mathrm{t}(x,y) = \int_{x-l_\mathrm{t}/2}^{x+l_\mathrm{t}/2} \int_{y-w_\mathrm{t}/2}^{y+w_\mathrm{t}/2} \frac{d(u,v)}{l_\mathrm{t} w_\mathrm{t}} \mathrm{d}u \mathrm{d}v, \tag{29}$$

where $d_0 = \sqrt{(x_\mathrm{a} - x_0)^2 + (y_\mathrm{a} - y_0)^2}$ is the distance between the antenna and the pixel, and the length and the width of the planar target are denoted by l_t and w_t.

The phase error $e_{2,\mathrm{p}}$ caused by the proposed method is expressed as

$$e_{2,\mathrm{p}} = \int_{x_0-l_\mathrm{s}/2}^{x_0+l_\mathrm{s}/2} \int_{y_0-w_\mathrm{s}/2}^{y_0+w_\mathrm{s}/2} \frac{2\pi |d_\mathrm{t}(x,y) - d_\mathrm{p}|}{l_\mathrm{s} w_\mathrm{s} \lambda} \mathrm{d}x \mathrm{d}y, \tag{30}$$

$$d_\mathrm{p} = \int_{x_0-l_\mathrm{s}/2}^{x_0+l_\mathrm{s}/2} \int_{y_0-w_\mathrm{s}/2}^{y_0+w_\mathrm{s}/2} \frac{d(\tilde{u},\tilde{v})}{l_\mathrm{s} w_\mathrm{s}} \mathrm{d}\tilde{u} \mathrm{d}\tilde{v}. \tag{31}$$

For planar targets, there are two cases: when the planar target does not completely fill the pixel, and when it fully fills the pixel. Incompletely filled pixels are usually at the edges of planar objects, while fully filled pixels are usually within planar objects. The error in both cases can be calculated using (30). The proposed method effectively cancels the phase errors for planar targets. In (28), there is a certain difference between d_0 calculated by the conventional algorithm and d_t in the planar target model. The d_p in (30) of the proposed model is close to d_t and makes the error as small as possible, especially when the planar target completely fills the pixel.

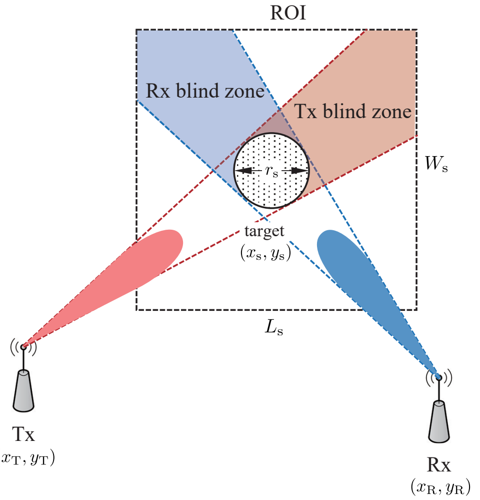


Fig. 7. Sensing ranges under different views.

B. Multi-View Sensing Accuracy

In this section, we analyze the accuracy of the sensing. At first, the sensing range differences between different transceiver views are crucial to the sensing performance. In addition, we analyze the impact of the sparse reconstruction sub-problem in (19) and the occlusion detection sub-problem in (22) on the algorithm performance. We use the mean square error (MSE) between the estimated value $\hat{\boldsymbol{x}}$ and the true value $\boldsymbol{x}$ of the environmental information to describe the accuracy of the sensing result as $\mathrm{MSE} = \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|_2^2 / N_\mathrm{s}$.

1) Multi-view Sensing Range Difference: In the model of (10), the sensing range difference caused by occlusion effect is reflected in the number of zero elements in the occlusion matrix $\mathbf{V}$. The proposed AO algorithm is to detect and eliminate occlusion effect as much as possible. In Fig. 7, we show the multi-view sensing scenario of a single transceiver pair as in (10). For an ideal circular target object with a random radius of $r_\mathrm{s} \in (0, R_\mathrm{s}]$ in ROI, there is a red blind zone for Tx and a blue blind zone for Rx. Let the positions of Tx and Rx be $\boldsymbol{p}^\mathrm{Tx} = [x_\mathrm{T}, y_\mathrm{T}]$, and $\boldsymbol{p}^\mathrm{Rx} = [x_\mathrm{R}, y_\mathrm{R}]$, respectively. The random target position is $\boldsymbol{p}^\mathrm{s} = [x_\mathrm{s}, y_\mathrm{s}]$ and position of the ROI center point is $\boldsymbol{p}^\mathrm{ROI} = [x_\mathrm{ROI}, y_\mathrm{ROI}]$. The occlusion matrix $\mathbf{V}$ is assumed to follow a binomial distribution.

Theorem 1: The probability that the element in the occlusion matrix $\mathbf{V}$ is 0 can be approximated as

$$\begin{aligned} p(\mathbf{V}_{n_\mathrm{s}} = 0) &= 1 - \lim_{N_\mathrm{s} \to \infty} \frac{\|\mathbf{V}\|_1}{N_\mathrm{s}^2} \\ &\approx \iint_{r_\mathrm{s}, \boldsymbol{p}^\mathrm{s}} \frac{r_\mathrm{s}\left((\|\boldsymbol{p}^\mathrm{Tx} - \boldsymbol{p}^\mathrm{ROI}\| + L_\mathrm{s}/2)_2^2 - \|\boldsymbol{p}^\mathrm{Tx} - \boldsymbol{p}^\mathrm{s}\|_2^2\right)}{2R_\mathrm{s} L_\mathrm{s} W_\mathrm{s} \|\boldsymbol{p}^\mathrm{Tx} - \boldsymbol{p}^\mathrm{s}\|_2} \\ &+ \iint_{r_\mathrm{s}, \boldsymbol{p}^\mathrm{s}} \frac{r_\mathrm{s}\left((\|\boldsymbol{p}^\mathrm{Rx} - \boldsymbol{p}^\mathrm{ROI}\| + L_\mathrm{s}/2)_2^2 - \|\boldsymbol{p}^\mathrm{Rx} - \boldsymbol{p}^\mathrm{s}\|_2^2\right)}{2R_\mathrm{s} L_\mathrm{s} W_\mathrm{s} \|\boldsymbol{p}^\mathrm{Rx} - \boldsymbol{p}^\mathrm{s}\|_2} \end{aligned} \tag{32}$$

where L_s and W_s are the side length of the square ROI.

Proof: The blind zone of a view can be approximated as an annular sector, whose inner and outer radii are the distance from the target to the antenna and the distance from the ROI edge to the antenna, respectively. Then considering the positions of antennas, and the random positions and sizes of the targets within the ROI, we can easily obtain (32). ■

Remark 1: It can be concluded that as the density of objects in the environment increases, the occlusion effect becomes more pronounced, leading to a higher number of zero elements in the occlusion matrix $\mathbf{V}$. This will increase the sensing error in the initial stage of iteration, resulting in errors in the occlusion detection, which will deteriorate the performance of the proposed AO algorithm.

2) Sparse Reconstruction Sub-problem: In this sub-problem, we aim to achieve the theoretical bound of CS sparse reconstruction after accomplishing optimal occlusion detection.

Theorem 2: The theoretical upper bound of the optimization problem in (19) under the optimal occlusion detection condition can be expressed as

$$\mathrm{MSE} \leq \frac{m(\log N_\mathrm{s})\|\boldsymbol{x}\|_1^2}{N_\mathrm{s}N_\mathrm{f}N_\mathrm{T}N_\mathrm{R}}, \tag{33}$$

where m is a constant.

Proof: Plug in the parameters in [21] or [4], a simple calculation yields (33). ■

Remark 2: It can be concluded that when the sensing overhead, such as the number of transmitting and receiving antennas and the number of carriers, increases, the sensing accuracy will increase. On the contrary, when the number of divided pixels increases and/or the environment objects become denser, the sensing accuracy will decrease.

3) Occlusion Detection Sub-problem: In this sub-problem, we focus on analyzing the impact of the occlusion matrix on the sensing performance. In the case of an unknown occlusion relationship, as shown in the model in (10) and (14), occlusion matrices $\mathbf{V}$ cause unknown interference to the compressed sensing measurement matrix $\mathbf{A}$. Especially in the early iteration stage of the proposed AO algorithm, when the estimated occlusion matrix $\hat{\boldsymbol{v}}^{\mathrm{UE}n_\mathrm{T},(t)}$ in (22) is not accurate, it affects the sensing performance more.

Theorem 3: For given ground truth, the Cramer-Rao bound (CRB) [31] of estimating the environmental information x_{n_s} under unknown occlusion interference in the proposed AO algorithm, is given as

$$\mathrm{MSE} \geq -E\left[\frac{\partial^2 \ln p(\hat{\mathbf{H}}; x_{n_\mathrm{s}})}{\partial (x_{n_\mathrm{s}})^2}\right]^{-1}, \tag{34}$$

where $E(\cdot)$ is mathematic expectation, and $p(\hat{\mathbf{H}}; x_{n_\mathrm{s}})$ is the likelihood function of x_{n_s} and $\hat{\mathbf{H}}$, which can be calculated by numerical methods when the distribution of occlusion matrix is given.

Remark 3: The proof follows the same line as in [31], thus is omitted due to limited space. As the proposed AO algorithm is iterated, the unknown occlusion interference is also reduced. The analysis of the system performance during the iteration process will be left to future work, and extensive simulation results have verified the effectiveness and convergence of the proposed algorithm.

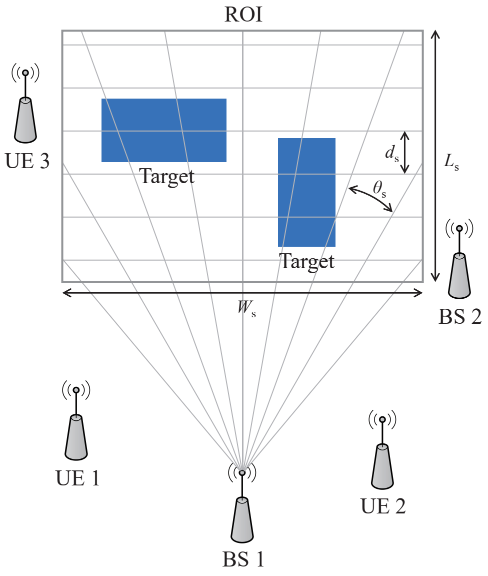


Fig. 8. The simulation scenario setting.

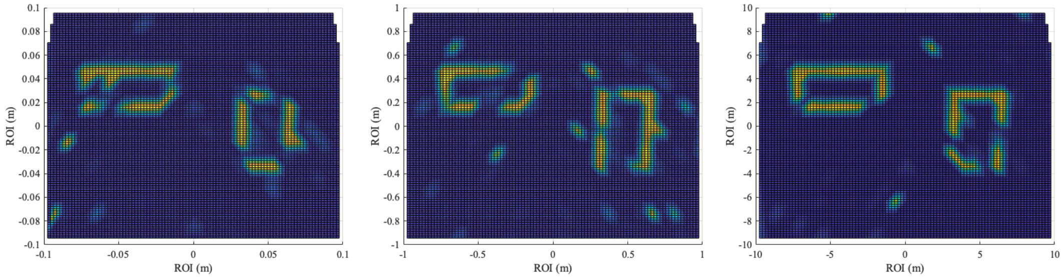
VI. NUMERICAL RESULTS

A. Simulation Scenario and Metrics

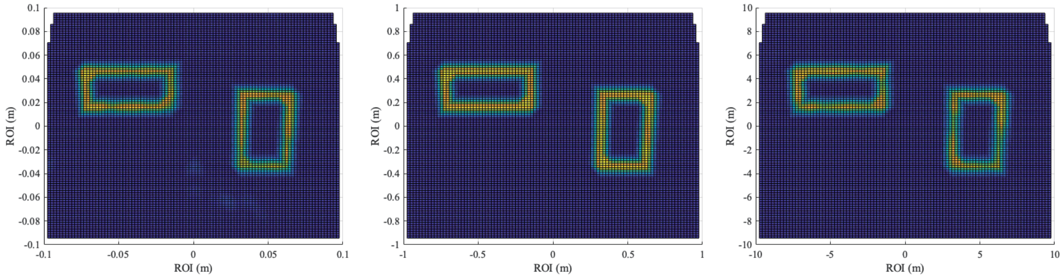
In this section, we present the numerical simulation results of the proposed method and the following algorithms are adopted as baselines: the OMP algorithm [32], a widely recognized CS reconstruction method with low computational complexity; the GAMP algorithm [23], which is a component of our proposed method but used here without considering occlusion effects; and the GAMP-MVSVR algorithm [8], an advanced centralized multi-view sensing method that achieves high reconstruction accuracy but lacks deployment flexibility and incurs significantly higher computational costs. In the GAMP-MVSVR algorithm, a central server collects and processes all channel gains between the UEs and BSs to perform global environment sensing, with the pixel grid defined from the view of BS1. All these baselines are implemented under the same propagation model introduced in Section III-A, ensuring fair comparison and applicability to large-scale scenarios.[12] We set up a square 2D ROI with targets deployed in it. For targets within the ROI, we employ the enhanced ray-tracing model proposed in [33] to obtain the received signals and propagation gains. Since the main idea of this paper is not multi-carrier sensing, the OFDM signaling parameters used for sensing are summarized as follows: a center frequency of 30 GHz, subcarrier spacing $\Delta f = 120$ kHz, and $N_\mathrm{f} = 2$ subcarriers spaced 100 MHz apart which can be dynamic in larger scale scenarios. Missed detection (MD) and false alarm (FA) rates are used to measure detection accuracy by

---

[12] It is worth noting that, the baseline GAMP-MVSVR used in our simulations adopts the same propagation model proposed in Section III-A, which actually enhances its original performance. Since this centralized algorithm performs global processing without mapping errors, it can be regarded as the performance upper bound of the proposed distributed algorithm.

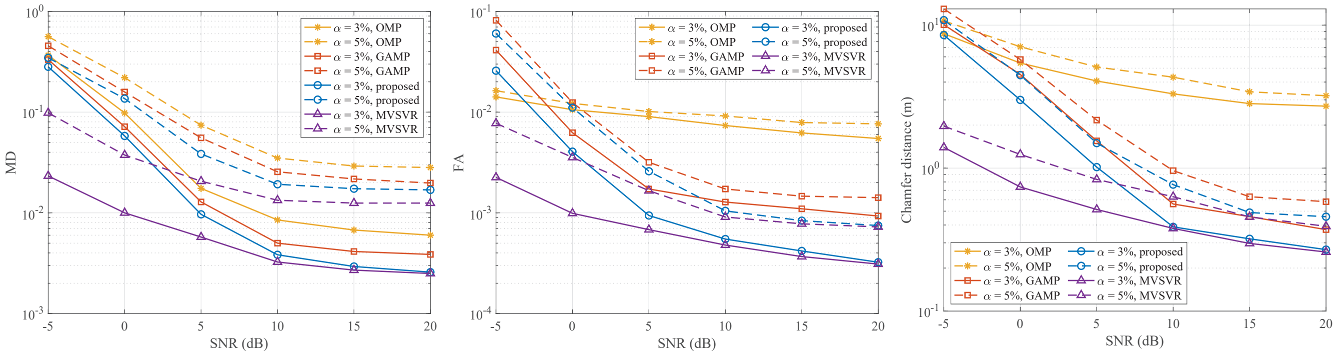


(a) Pixel interval $d_{\mathrm{s}} = 1$ cm, the baseline method sensing result.
(b) Pixel interval $d_{\mathrm{s}} = 10$ cm, the baseline method sensing result.
(c) Pixel interval $d_{\mathrm{s}} = 1$ m, the baseline method sensing result.



(d) Pixel interval $d_{\mathrm{s}} = 1$ cm, the proposed method sensing result.
(e) Pixel interval $d_{\mathrm{s}} = 10$ cm, the proposed method sensing result.
(f) Pixel interval $d_{\mathrm{s}} = 1$ m, the proposed method sensing result.

Fig. 9. The multi-view sensing results of the baseline method and proposed method.



(a) The relationship between the SNR and MD.
(b) The relationship between the SNR and FA.
(c) The relationship between the SNR and CD.

Fig. 10. The relationship between the SNR and multi-view sensing performance.

comparing the ground truth object $\boldsymbol{x}$ with the estimated results $\hat{\boldsymbol{x}}$,

$$\mathrm{MD} = \frac{\sum_{n_{\mathrm{s}}=1}^{N_{\mathrm{s}}} \mathbb{I}\{x_{n_{\mathrm{s}}} \neq 0, \hat{x}_{n_{\mathrm{s}}} < \rho\}}{\sum_{n_{\mathrm{s}}=1}^{N_{\mathrm{s}}} \mathbb{I}\{x_{n_{\mathrm{s}}} \neq 0\}}, \tag{35}$$

$$\mathrm{FA} = \frac{\sum_{n_{\mathrm{s}}=1}^{N_{\mathrm{s}}} \mathbb{I}\{x_{n_{\mathrm{s}}} = 0, \hat{x}_{n_{\mathrm{s}}} > \rho\}}{\sum_{n_{\mathrm{s}}=1}^{N_{\mathrm{s}}} \mathbb{I}\{x_{n_{\mathrm{s}}} = 0\}}, \tag{36}$$

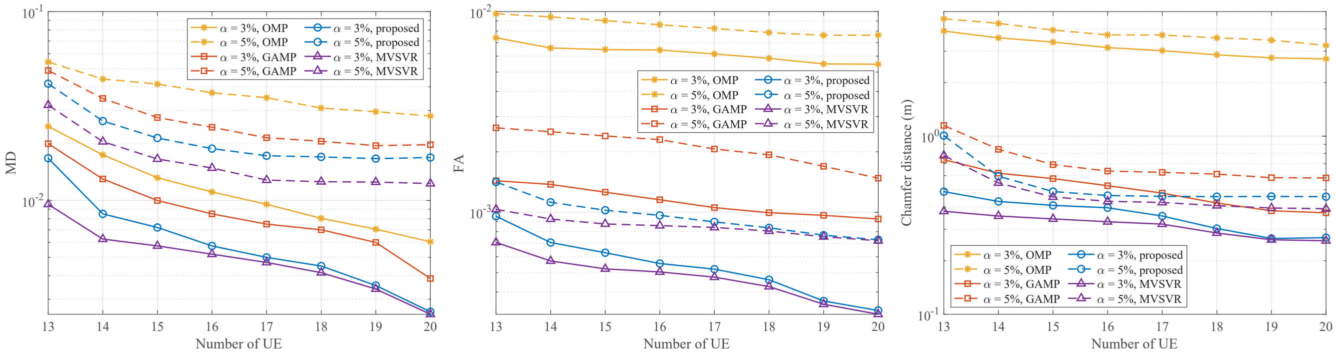
where where $\mathbb{I}(\cdot)$ is the indicator function and we normalize the scatterer coefficient and set the threshold $\rho = 0.5$ to determine whether a pixel contains a scatterer. A smaller MD and FA indicate fewer missed targets and false detections, thus reflecting better detection performance. Chamfer distance (CD) is used to measure the spatial accuracy of the sensing results,

$$d_{\mathrm{CD}}(\mathcal{X}, \hat{\mathcal{X}}) = \frac{1}{|\mathcal{X}|} \sum_{\mathbf{p}(x)\in\mathcal{X}} \min_{\mathbf{p}(\hat{x})\in\hat{\mathcal{X}}} \|\mathbf{p}(x) - \mathbf{p}(\hat{x})\|_2^2 + \frac{1}{|\hat{\mathcal{X}}|} \sum_{\mathbf{p}(\hat{x})\in\hat{\mathcal{X}}} \min_{\mathbf{p}(x)\in\mathcal{X}} \|\mathbf{p}(\hat{x}) - \mathbf{p}(x)\|_2^2, \tag{37}$$

where $\mathcal{X}$ and $\hat{\mathcal{X}}$ are pixel sets of the ground truth and estimation and $\mathbf{p}(\cdot)$ denotes the spatial coordinate of the pixel. A smaller d_{CD} indicates a closer match between the two pixel sets.

As shown in Fig. 8, we deploy rectangular targets in the ROI as a simulated urban target setting for intuitive imaging results. The BSs and UEs are randomly and uniformly distributed within an area ten times larger than the ROI, with the ROI located at the center, and we use the coordinate system division under the view of BS1 to show the imaging results. In addition, all other performance simulations in this paper are conducted based on randomly generated planar targets, which simulate the urban building wall structure within the ROI.

Fig. 9 shows the intuitive results of both the baseline and the proposed methods for multi-view imaging of rectangular

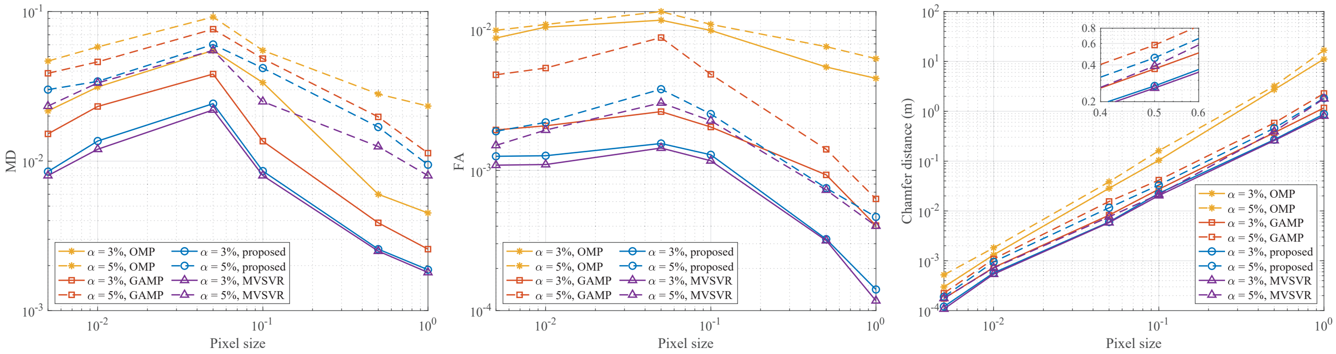


(a) The relationship between the number of UE and MD.

(b) The relationship between the number of UE and FA.

(c) The relationship between the number of UE and CD.

Fig. 11. The relationship between the number of UE and multi-view sensing performance.

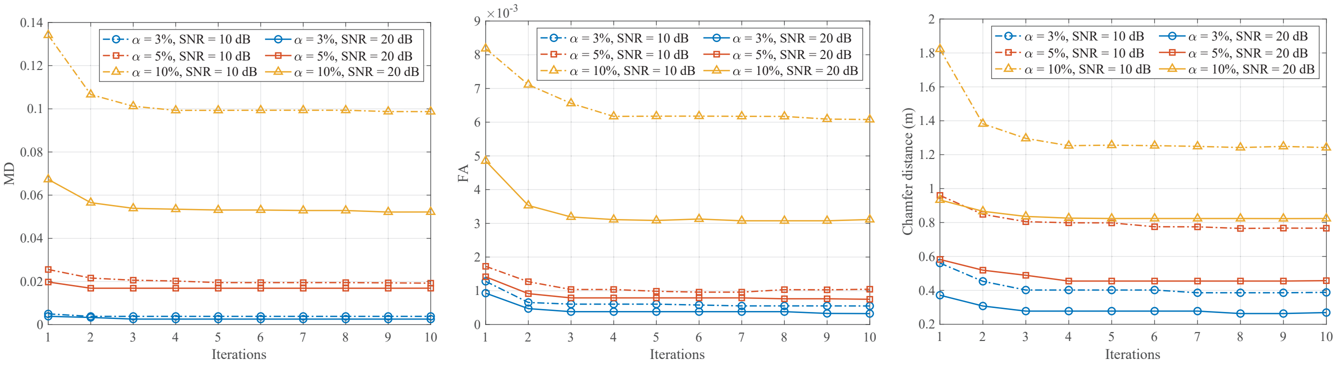


(a) The relationship between the pixel size and MD.

(b) The relationship between the pixel size and FA.

(c) The relationship between the pixel size and CD.

Fig. 12. The relationship between the pixel size and multi-view sensing performance.



(a) The convergence performance in MD.

(b) The convergence performance in FA.

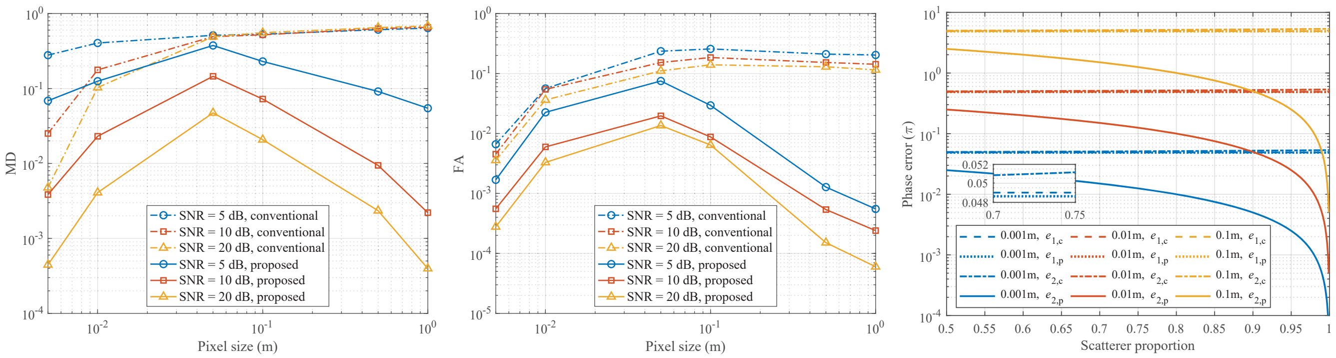
(c) The convergence performance in CD.

Fig. 13. The convergence performance of multi-view sensing.

urban buildings. The color of the pixel represents the scattering coefficient of the corresponding pixel. There are $N_{\mathrm{T}} = N_{\mathrm{R}} = 15$ transceivers deployed around the square ROI shown in Fig. 9. Each transceiver has its own coordinate system for the corresponding view. The division intervals of the non-uniform coordinate system are $1\mathrm{cm}, 10\mathrm{cm}, 1\mathrm{m}$ and the division angles are $0.05°, 0.1°, 0.2°$. Compared with the traditional computational imaging algorithm described in Section IV-B that ignores the concept of multi-view, the proposed sensing method obtains the shape of the target more accurately. At the same time, benefiting from the proposed pixel division error cancellation method, we achieved environment sensing under different pixel sizes.

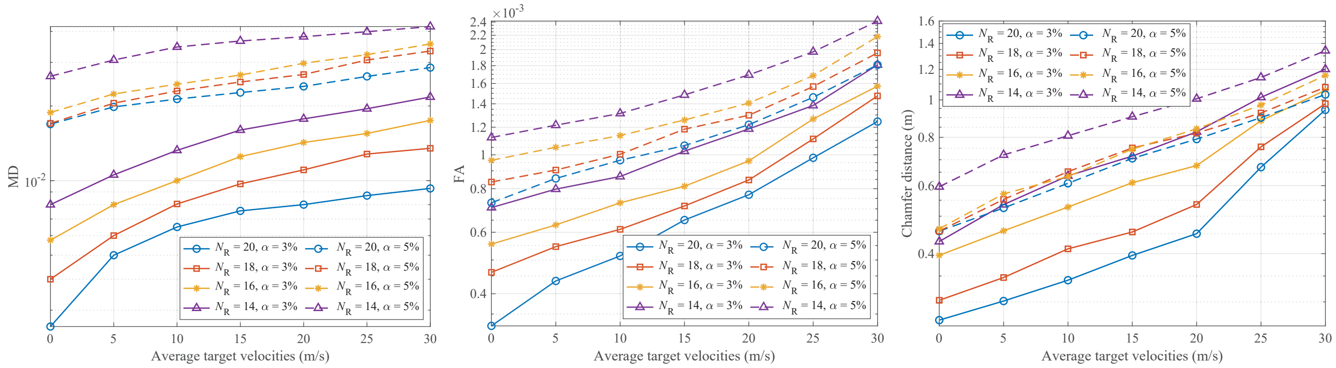
B. *Performance Evaluation*

Fig. 10 shows the multi-view sensing performance of the baseline and the proposed algorithms under different signal-to-noise ratios (SNRs). SNR is defined as the power ratio of the channel estimation result to the noise in (14). The number of transceivers is set to $N_{\mathrm{T}} = N_{\mathrm{R}} = 20$. The division intervals of the non-uniform coordinate system are 0.5 m. It can be observed that the sensing performance improves as the SNR increases because higher SNR leads to smaller



(a) The relationship between the pixel size and MD. (b) The relationship between the pixel size and FA. (c) The relationship between the scatterer proportion and phase error under different pixel sizes.

Fig. 14. The relationship between the pixel size and environment sensing performance.



(a) The relationship between the average target velocities and MD. (b) The relationship between the average target velocities and FA. (c) The relationship between the average target velocities and CD.

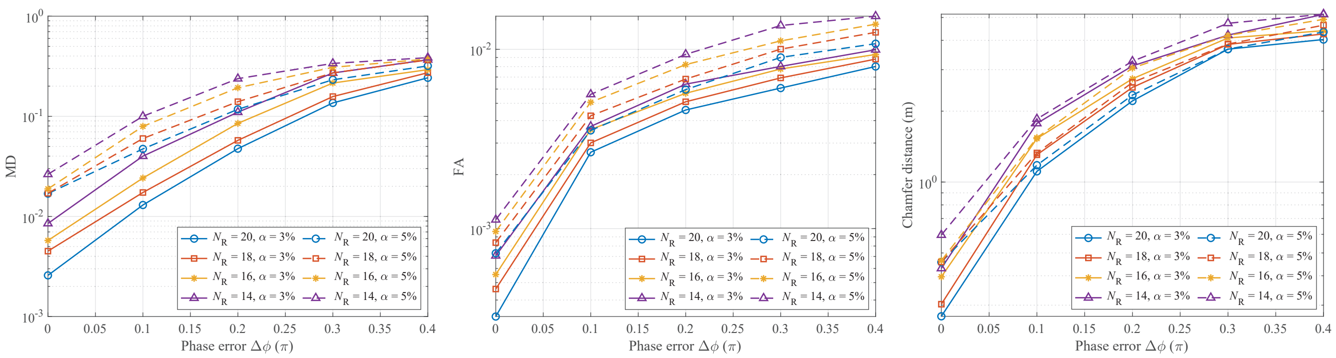
Fig. 15. The relationship between the average target velocities and multi-view sensing performance.

channel estimation errors, which in turn enhances the accuracy of reconstruction. Furthermore, as analyzed in Theorem 2, a lower sparsity also facilitates a more accurate reconstruction. In addition, as the number of targets increases, accurately identifying occlusion relationships becomes more difficult, thus reducing imaging accuracy. The proposed algorithm consistently outperforms the GAMP algorithm. Due to the pursuit mechanism, OMP maintains a relatively low FA rate under low SNR conditions, however, its overall performance remains inferior to the proposed method. The proposed distributed algorithm achieves performance comparable to the centralized GAMP-MVSVR algorithm under high SNR conditions. Its performance degradation at low SNR is mainly attributed to the mismatch of occlusion matrices during multi-view mapping, which becomes more pronounced as noise increases.

Fig. 11 shows the multi-view sensing performance of the baseline algorithm and the proposed algorithm under different numbers of UEs (N_{T}), and the number of BSs remains constant ($N_{\mathrm{R}} = 20$). The SNR is set to 20 dB. The division intervals of the non-uniform coordinate system are 0.5 m. The performance improves as the number of UEs increases. This is because a larger number of UEs provides more observations of the environment, which enriches the available information and improves the accuracy of sensing. This phenomenon is also consistent with the analysis in Theorem 2. The proposed algorithm consistently outperforms the OMP and GAMP algorithms and approaches the performance of the centralized GAMP-MVSVR algorithm as the number of users increases.

Fig. 12 shows the multi-view sensing performance of the baseline algorithm and the proposed algorithm under different pixel sizes. The number of transceivers is set to $N_{\mathrm{T}} = N_{\mathrm{R}} = 20$ and SNR $= 20$ dB. The proposed algorithm consistently outperforms the OMP and GAMP algorithms and approaches the performance of the centralized GAMP-MVSVR algorithm. When the pixel size is small (comparable to the wavelength), increasing the pixel size introduces additional phase errors, which become the dominant factor causing performance degradation. When the pixel size is further enlarged, conventional computational imaging fails completely (as verified in Fig. 14), while the proposed method still perform accurate reconstructions. After eliminating the phase error, according to the principles of computational imaging, larger pixel sizes reduce the correlation of the CS measurement matrix, which improves reconstruction performance and decreases MD and FA. Meanwhile, the CD grows approximately linearly with pixel size.

Fig. 13 shows the convergence curves of MD, FA and CD in Fig. 10 when SNRs are 10 dB and 20 dB. We can observe that with the iteration of the proposed AO algorithm, the system performance gradually converges. In Fig. 13, we increase the



(a) The relationship between the phase synchronization error and MD.

(b) The relationship between the phase synchronization error and FA.

(c) The relationship between the phase synchronization error and CD.

Fig. 16. The multi-view sensing performance under different sparsity levels α when phase synchronization errors are introduced between UEs and BSs.

number of iterations to prove that the algorithm reaches a stable convergence. The above figures also show the relationship between environment sparsity and system performance. As shown in Theorems 1 and 3, an increased number of objects leads to more occlusions, introducing more unknown entries in the occlusion matrix $\mathbf{V}$. This makes occlusion detection and sparse recovery more difficult, thereby degrading performance. In addition, the algorithm converges more quickly under higher SNR because the initial imaging results are more accurate, which accelerates subsequent iterations and improves convergence behavior.

Fig. 14(a) and (b) shows the environment sensing results under different pixel sizes, and we focus on verifying the effectiveness of the pixel division error cancellation method proposed in Sections III-A and V-A. The system parameters are set to $N_{\mathrm{T}} = N_{\mathrm{R}} = 20$, the number of uniform pixels is $N_{\mathrm{s}} = 30 \times 30$, and SNR $= 20$ dB. The results provide an independent validation of the conclusions in Fig. 12 regarding pixel size. The same trend can be observed: as pixel size increases, system performance first improves and then degrades. This is the result of the joint effects of pixel division errors and the spatial correlation induced by large pixels. Fig. 14(c) shows the numerical integration results for the phase error of the point targets and the planar targets in Section V-A. The proposed method significantly cancels the phase errors. Especially in the case of planar targets, as the proportion of scatterers $(l_{\mathrm{t}}w_{\mathrm{t}})/(l_{\mathrm{s}}w_{\mathrm{s}})$ within a pixel increases, the phase error decreases, enabling accurate computational imaging. In practical environments, strictly planar targets that partially fill a pixel occur only at limited boundary regions, and ideal point targets rarely exist. Even in the most unfavorable cases, the proposed model reduces to the conventional point-based model and thus achieves at least the same modeling performance as conventional methods.

C. Robustness Evaluation

In this section, we provide a robustness evaluation to demonstrate the performance of the proposed method under various interferences. Denote the projection of the vehicle's velocity in the angle-of-arrival direction as v, then the resulting Doppler shift is expressed as $e^{\jmath 2\pi T v/\lambda}$, where T denotes the transmission duration of each symbol. Let the transmission duration of each symbol be $T = 1/30$ kHz, Fig. 15 presents the impact of Doppler effects on system performance for average target velocities v under different sparsity levels α. It can be concluded that as the vehicle speed increases, Doppler weakens the system performance more significantly, but the proposed algorithm remains effective. The proposed framework, like other coherent imaging methods, assumes a quasi-static scene during each coherent processing interval. This assumption does not preclude operation in dynamic scenarios, as motion-induced phase variations can be controlled by appropriately selecting system parameters. Explicit motion modeling and tracking can be incorporated as extensions and are left for future work.

Fig. 16 shows the multi-view sensing performance under different sparsity levels α when phase synchronization errors are introduced between UEs and BSs. The synchronization errors are modeled as a Gaussian random variable $\mathcal{N}(0, \Delta\phi)$, which may result from inaccurate UE positioning (whose error affects each row of the measurement matrix and is equivalent to the phase synchronization error) or hardware imperfections. The results show that phase synchronization errors degrade sensing performance, which is a fundamental limitation shared by coherent imaging methods [34]. When the synchronization error is relatively small, the proposed algorithm remains effective. In the quasi-static sensing scenarios considered in this paper, near-nanosecond synchronization accuracy can be achieved by (i) system-level high-precision timing (e.g., GNSS-based timing), (ii) network-level multi-node collaborative synchronization techniques, and (iii) signal-level or even sample-level time-frequency synchronization. In addition, the large array gain of massive antenna systems in general leads to high sensitivity to temporal and spatial phase variations and thus results in finer resolution in time and phase calibration as in our system. These measures can be carried out in parallel to refine and calibrate the system synchronization, which is expected to meet the synchronization accuracy requirement for the studied cases. Detailed implementation and experimental validation are left for future work.

VII. CONCLUSION

In this paper, we have proposed a distributed multi-view sensing architecture with a unique pixel coordinate system for each transceiver, based on non-uniform pixel division and mapping. We have proposed a signal propagation model to cancel pixel division phase errors and overcome occlusion effects, enabling large-scale environment sensing. A low-complexity alternating optimization algorithm is developed, combining CS sparse reconstruction and occlusion detection to achieve accurate sensing results. This approach offers insights into the ISAC method's application in large-scale scenarios.

REFERENCES

[1] W. Saad et al., "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.

[2] F. Liu et al., "Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

[3] Z. Wei et al., "Integrated sensing and communication signals toward 5G-A and 6G: A survey," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11068–11092, 2023.

[4] X. Tong et al., "Joint multi-user communication and sensing exploiting both signal and environment sparsity," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1409–1422, 2021.

[5] Y. Cui and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Network*, vol. 35, no. 5, pp. 158–167, 2021.

[6] V. Degli-Esposti et al., "Measurement and modelling of scattering from buildings," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 1, pp. 143–153, 2007.

[7] Z. Wei et al., "Integrated sensing and communication enabled multiple base stations cooperative sensing towards 6G," *IEEE Network*, vol. 38, no. 4, pp. 207–215, 2024.

[8] X. Tong et al., "Environment sensing considering the occlusion effect: A multi-view approach," *IEEE Transactions on Signal Processing*, vol. 70, pp. 3598–3615, 2022.

[9] E. C. Strinati et al., "Distributed intelligent integrated sensing and communications: The 6G-DISAC approach," in *EuCNC/6G Summit*, 2024, pp. 392–397.

[10] X. Tong et al., "Multi-view sensing for wireless communications: Architectures, designs, and opportunities," *IEEE Communications Magazine*, vol. 61, no. 5, pp. 40–46, 2023.

[11] G. Li et al., "Multi-point integrated sensing and communication: Fusion model and functionality selection," *IEEE Wireless Communications Letters*, vol. 11, no. 12, pp. 2660–2664, 2022.

[12] W. Jiang et al., "Collaborative precoding design for adjacent integrated sensing and communication base stations," *IEEE Internet of Things Journal*, vol. 11, no. 9, pp. 15059–15074, 2024.

[13] P. Liu et al., "Vertical federated edge learning with distributed integrated sensing and communication," *IEEE Communications Letters*, vol. 26, no. 9, pp. 2091–2095, 2022.

[14] D. Tagliaferri et al., "Cooperative coherent multistatic imaging and phase synchronization in networked sensing," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 10, pp. 2905–2921, 2024.

[15] M. Manzoni et al., "Wavefield networked sensing: Principles, algorithms, and applications," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 181–197, 2025.

[16] A. Liu et al., "A survey on fundamental limits of integrated sensing and communication," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 994–1034, 2022.

[17] Y. Ge et al., "A computationally efficient EK-PMBM filter for bistatic mmwave radio SLAM," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2179–2192, 2022.

[18] Z. Tian et al., "Distributed multi-view sparse vector recovery," *IEEE Transactions on Signal Processing*, vol. 71, pp. 1448–1463, 2023.

[19] B. Lu et al., "Deep learning based multi-node ISAC 4D environmental reconstruction with uplink-downlink cooperation," *IEEE Internet of Things Journal*, pp. 1–1, 2024.

[20] B. Sun et al., "3D computational imaging with single-pixel detectors," *Science*, vol. 340, no. 6134, pp. 844–847, 2013.

[21] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[22] T. T. Cai et al., "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.

[23] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *2011 IEEE ISIT*, 2011, pp. 2168–2172.

[24] N. Vukmirović et al., "Direct wideband coherent localization by distributed antenna arrays," *Sensors*, vol. 19, no. 20, pp. 4582, 2019.

[25] Y. Zhang et al., "3D environment sensing with channel state information based on computational imaging," in *2022 IEEE ICC Workshops*, 2022, pp. 842–847.

[26] Y. Niu et al., "A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges," *Wireless networks*, vol. 21, pp. 2657–2676, 2015.

[27] M. A. Richards et al., *Fundamentals of radar signal processing*, Mcgraw-hill, 2005.

[28] Q. Wu et al., "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.

[29] Z. Lin et al., "A compressive sensing approach for MIMO-OFDM-based integrated sensing and communication," in *2023 IEEE GLOBECOM*, 2023, pp. 522–527.

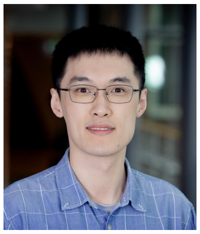
[30] X. Tong et al., "Environment sensing with beam sweeping and non-uniform pixelation in wireless communication systems," in *IEEE PIMRC*, 2023, pp. 1–7.

[31] X. Song et al., "Intelligent reflecting surface enabled sensing: Cramér-Rao bound optimization," *IEEE Transactions on Signal Processing*, vol. 71, pp. 2011–2026, 2023.

[32] J. A. Tropp et al., "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[33] Z. Xing et al., "Physics-inspired target shape detection and reconstruction in mmWave communication systems," in *IEEE GLOBECOM 2023*, 2023, pp. 3910–3915.

[34] J. Pegoraro et al., "MovISAC: Coherent imaging of moving targets with distributed asynchronous ISAC devices," *arXiv preprint arXiv:2502.08236*, 2025.

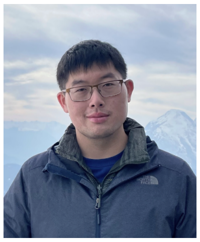


Xin Tong (Member, IEEE) received his B.S. degree from Ningbo University, Ningbo, China, in 2020 and Ph.D. degree from Zhejiang University, Hangzhou, China, in 2025, under the supervision of Prof. Zhaoyang Zhang. He is currently with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, and with Zhejiang Provincial Laboratory of Multi-Modal Communication Networks and Intelligent Information Processing, Hangzhou 310027, China, and also with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden. His research interests include integrated sensing and communication, signal processing, and distributed MIMO.

Zhaoyang Zhang (Senior Member, IEEE) received his Ph.D. degree from Zhejiang University, Hangzhou, China, in 1998, where he is currently a Qiushi Distinguished Professor. His research interests are mainly focused on the fundamental aspects of wireless communications and networking, with emphases on AI-empowered communications and networking, integrated communication, sensing and computing, and AI sensing, etc. He has co-authored more than 200 IEEE journal papers, and is a co-recipient of 2024 IEEE Leonard G. Abraham Prize and over 10 best paper awards, best student paper awards, or student travel award of IEEE ICC 2019, Globecom 2020, ISIT 2023, WCNC 2024, and PIMRC 2025, etc. He was granted the National Natural Science Fund for Distinguished Young Scholar by NSFC in 2017, and was a co-recipient of the First Grade State-level Teaching Award for Graduate Education in 2023.

Dr. Zhang served as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE TRANSACTIONS ON COMMUNICATIONS, etc., the Lead Guest Editor for IEEE WIRELESS COMMUNICATIONS Special Issue on Sustainable Big AI Model for Wireless Networks, the General Co-Chair or TPC Co-Chair for WCSP 2023/2018/2013, PIMRC 2021 Workshop on Native AI Empowered Wireless Networks, and VTC-Spring 2017 Workshop on HMWC, etc. He was also a keynote speaker of IEEE Globecom 2021 Workshop on Native-AI, APCC 2018 and VTC-Fall 2017 Workshop on NOMA, etc.

Zhaohui Yang (Member, IEEE) is currently a ZJU Young Professor with the Zhejiang Key Laboratory of Information Processing Communication and Networking, College of Information Science and Electronic Engineering, Zhejiang University. He received the Bachelor and Ph.D. degrees from Southeast University, Nanjing, China, in 2014 and 2018, respectively. From 2018 to 2020, he was a Post-Doctoral Research Associate at the Center for Telecommunications Research, Department of Informatics, King's College London, U.K. From 2020 to 2022, he was a Research Fellow at the Department of Electronic and Electrical Engineering, University College London, U.K. His research interests include joint communication, sensing, and computation, federated learning, and semantic communication. He received IEEE Communications Society Asia-Pacific Outstanding Young Researcher Award, IEEE Marconi Prize Paper Award, and IEEE Katherine Johnson Young Author Paper Award. He currently serves as an Associate Editor of IEEE TGCN, IEEE CL, IEEE TMLCN. He has served as a Guest Editor for several journals including IEEE Journal on Selected Areas in Communications.



Yu Ge (Member, IEEE) received his B.E. degree from Zhejiang University, Hangzhou, China, in 2017, M.Sc. degree from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2019, and Ph.D. degree from Chalmers University of Technology, Sweden, in 2024. He was a Postdoctoral Researcher with the Department of Electrical Engineering at Chalmers University of Technology, Sweden. His research interests include integrated communication and sensing, wireless positioning systems, simultaneous localization and mapping, and multi-object tracking, particularly in 5G and Beyond 5G scenarios.

Henk Wymeersch (Fellow, IEEE) obtained the Ph.D. degree in Electrical Engineering/Applied Sciences in 2005 from Ghent University, Belgium. He is currently a Professor of Communication Systems with the Department of Electrical Engineering at Chalmers University of Technology, Sweden and a Distinguished Visiting Professor at Tsinghua University. Prior to joining Chalmers, he was a postdoctoral researcher from 2005 until 2009 with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. Prof. Wymeersch served as Associate Editor for IEEE Communication Letters, IEEE Transactions on Wireless Communications, and IEEE Transactions on Communications and is currently Senior Member of the IEEE Signal Processing Magazine Editorial Board. During 2019-2021, he was an IEEE Distinguished Lecturer with the Vehicular Technology Society. His research interests include the convergence of communication and sensing, in a 5G and Beyond 5G context.