Enhancing Sensing-Assisted Communications in Cluttered Indoor Environments Through Background Subtraction

N.B. When citing this work, cite the original published paper.

(article starts on next page)

Enhancing Sensing-Assisted Communications in Cluttered Indoor Environments through Background Subtraction

Andrea Ramos*, Musa Furkan Keskin†, Henk Wymeersch†, Saúl Inca*, Jose F. Monserrat*
*iTEAM Research Institute, Universitat Politècnica de València, Spain
†Chalmers University of Technology, Sweden

Abstract—Integrated sensing and communications (ISAC) is poised to be a native technology for the forthcoming Sixth Generation (6G) era, with an emphasis on its potential to enhance communications performance through the integration of sensing information, i.e., sensing-assisted communications (SAC). Nevertheless, existing research on SAC has predominantly confined its focus to scenarios characterized by minimal clutter and obstructions, largely neglecting indoor environments, particularly those in industrial settings, where propagation channels involve high clutter density. To address this research gap, background subtraction is proposed on the monostatic sensing echoes, which effectively addresses clutter removal and facilitates detection and tracking of user equipments (UEs) in cluttered indoor environments with SAC. A realistic evaluation of the introduced SAC strategy is provided, using ray tracing (RT) data with the scenario layout following Third Generation Partnership Project (3GPP) indoor factory (InF) channel models. Simulation results show that the proposed approach enables precise predictive beamforming largely unaffected by clutter echoes, leading to significant improvements in effective data rate over the existing SAC benchmarks and exhibiting performance very close to the ideal case where perfect knowledge of UE location is available.

Index Terms—Indoor factory, ISAC, 6G, sensing-assisted communications, background subtraction.

I. INTRODUCTION

In recent years, wireless communication systems, especially Fifth Generation (5G), has played a crucial role in meeting the diverse and demanding requirements of the Fourth Industrial Revolution, commonly known as Industry 4.0 [1]. Industry 4.0 strives to create adaptable and efficient smart factories by integrating applications of industrial internet of Things (IIoT). These applications are used in various areas such as logistics, including motion control, smart transportation for inventory management, and collaborative intelligent robots for manufacturing [2]. These operations frequently occur in indoor scenarios with numerous machines and metallic surfaces, which might pose challenges for wireless connectivity.

Indoor factory (InF) environments can be more demanding than other indoor deployments [3]. The presence of obstacles results in multiple reflections, leading to a large number of Multipath Components (MPCs). Consequently, establishing high-directional communication may depend on accurately localizing the user's position [4], [5]. Moreover, localization can benefit from exploiting characteristics associated with frequency bands [6], which could be an additional feature using the same radio frequency (RF) resources.

Within the new era of Sixth Generation (6G), numerous studies exist on high-resolution localization and sensing [7], [8]. Following this burgeoning interest, novel technologies are poised to emerge, significantly enhancing communication systems through precise localization. One such groundbreaking concept is integrated sensing and communications (ISAC) systems [9]. ISAC is an enabling technology that thrives on the coexistence between sensing and communication capabilities. It is based on the fundamental premise that a sensing system can be seamlessly adapted to multiple radio technologies and diverse environments [10]. The convergence of these functionalities promises a host of benefits, including the ability for a communication system to act as a sensor or sensing-assisted communications (SAC) [11].

One of the earliest contributions of SAC is in [12], where sensing serves communication with novel predictive beamforming (beam tracking). This solution was introduced to mitigate the time overhead caused by downlink and uplink pilots in the conventional channel state information (CSI) process. The authors also proposed a simple method to handle beam association of multi-vehicle tracking. A parallel endeavor is evident in [13], which proposes an ID association technique to efficiently predict the state of multiple vehicles. This technique leverages the Kullback-Leibler divergence (KLD) to discern which ID corresponds to each vehicle without needing feedback at every time. Both contributions consider a reflected echo per vehicle for simplicity, an assumption that may not work in cluttered environments.

Another SAC contribution is introduced in [14], which proposes dynamic predictive beamforming. This concept involves the adaptation of the beamwidth to track extended vehicle. While this approach might fully illuminate the vehicle as needed, rather than a pencil-sharp beam, the authors leave the scenario geometry out of scope in terms of channel characterization in the system model. In research focused on
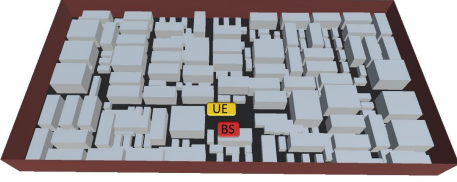
Fig. 1: InF sub-scenario with high clutter density (Dense High). The red square represents the base station (BS), and the yellow square represents the user equipment (UE).

ISAC systems, which depend on a thorough comprehension of the scenario, it is advisable to employ realistic assumptions in channel modeling, such as the inclusion of multipath channels. Although all aforementioned works have contributed to shed light on SAC systems, they have all employed an analytical channel model.

In terms of multipath channels, authors in [15] model a theoretical channel considering Line of Sight (LoS) and Non Line of Sight (NLoS) paths, focusing on the quantification of SAC overhead reductions. A more elaborate channel modeling is introduced in [16], which considers Third Generation Partnership Project (3GPP) geometry-based stochastic model (GBSM) for the communication channel and multiple-input multiple-output (MIMO) radar multipath channel for sensing one. Even though the latter contribution may be closer to realistic performance, spatial consistency and correlation between sensing and communication channels are inherent features that deterministic channel models (e.g., ray tracing (RT) dataset) may naturally include. Moreover, the previous contributions only focus on vehicle-to-anything (V2X) communication, where environments are "cleaner" with fewer obstructions. Most likely, in more clutter-dense scenarios, detecting the user among obstacles may be challenging, as in InF, as shown in Fig. 1. In industrial environments, where the reliability of communication is paramount for IIoT applications, the clutter in the scenario can directly impact the LoS communication. Thus, sensing could play a crucial role in precisely locating the user in such environments.

This paper proposes background subtraction as a possible method to facilitate user detection and improve predictive beamforming in SAC systems in cluttered indoor environments. This work aims to take a step towards a realistic evaluation of the ISAC system by obtaining realistic measurements from a RT tool. In order to reveal how beam misalignment can be affected by the number of passive scatterers in the InF environment, signal-to-noise ratio (SNR) and effective date rate are selected as key performance indicator (KPI). To establish a comparison, several benchmarks are considered, including SAC approaches and communication-only systems.

II. System Model and Performance Metrics

A downlink communications scenario is considered for an InF environment with a multiple-antenna BS and a single-antenna UE. The BS transmits data symbols to the UE while collecting the backscattered signals via a multiple-antenna co-

located radar receiver for detecting and tracking the UE to enable beam tracking leveraging the principles of SAC.

A. Sensing Signal Model

The BS operates with MIMO uniform linear array (ULA), which has N_T transmit antennas and N_R receive antennas. Given a precoder $\mathbf{f}$, the backscattered signal across N subcarriers can be formulated as:

$$\mathbf{y}_n^{\mathrm{sen}} = \mathbf{H}_n \mathbf{f} x_n + \mathbf{z}_n^s \in \mathbb{C}^{N_\mathrm{R} \times 1} \ , \tag{1}$$

where x_n is the Orthogonal frequency Division Multiplexing (OFDM) transmitted signal considering a transmitted power P; $\mathbf{z}_n^s$ denotes the additive white Gaussian noise (AWGN) with zero mean with a variance σ_N^2, and

$$\mathbf{H}_n = \sum_{k=1}^{K} \alpha_k^s e^{-j2\pi n \Delta f \tau_k^s} \mathbf{a}_\mathrm{R}(\theta_k^s) \mathbf{a}_\mathrm{T}^\top(\theta_k^s) \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{T}} \ , \tag{2}$$

where Δf denotes subcarrier spacing, α_k^s is the complex channel gain, θ_k^s is the angle-of-departure (AoD) (equal to the angle-of-arrival (AoA)), and τ_k^s is the delay. The transmit steering vector is the same as the receive steering vector given by the monostatic sensing configuration (i.e., transmitter (Tx) and receiver (Rx) are co-located on the same hardware). Thus, both steering vectors, namely $\mathbf{a}_\mathrm{T}(\theta_k^s)$ and $\mathbf{a}_\mathrm{R}(\theta_k^s)$ can be denoted as:

$$\mathbf{a}(\theta_k^s) = [1, e^{-j\pi \sin(\theta_k^s)}, \ldots, e^{-j\pi(N_\mathrm{T}-1)\sin(\theta_k^s)}]^\top, \tag{3}$$

considering half-wavelength antenna spacing.

B. Communication Signal Model

Assuming a multiple input single output (MISO) downlink communication, the received signal at the UE can be written as

$$y_n^{\mathrm{com}} = (\mathbf{h}_n^{\mathrm{com}})^\top \mathbf{f} x_n + z_n^c \in \mathbb{C} \ , \tag{4}$$

where

$$\mathbf{h}_n^{\mathrm{com}} = \sum_{k=1}^{\widetilde{K}} \alpha_k^c e^{-j2\pi n \Delta f \tau_k^c} \mathbf{a}_\mathrm{T}(\theta_k^c) \in \mathbb{C}^{N_\mathrm{T} \times 1}. \tag{5}$$

C. KPI Selection

The goal of this work is to design the communication precoders of the form

$$\mathbf{f} = \mathbf{a}_\mathrm{T}^*(\theta) \in \mathbb{C}^{N_\mathrm{T} \times 1}, \tag{6}$$

for a certain beamforming angle θ. The selection of the beamforming angle depends on the specific method. Discrete frames with duration T are considered, comprising beam training and data transmission. After beam training, a vector $\mathbf{f}$ is determined.

The SNR at the UE achieved after the beam training period can be written as

$$\mathrm{SNR}_{r,n} = \frac{P|(\mathbf{h}_n^{\mathrm{com}})^\top \mathbf{f}|^2}{\sigma_N^2} \ . \tag{7}$$

The effective data rate is then formulated as

$$R_{\mathrm{eff}} = \left(\frac{T-D}{T}\right) \sum_{n=1}^{N} \log_2(1 + \mathrm{SNR}_{r,n}), \tag{8}$$

where D is the time needed for beam training.

III. BEAM TRAINING METHODS

The UE moves over discrete time t, and at each time t, a downlink precoder should be designed. When possible, the time index is removed to lighten the notation.

A. Conventional Beam Training

Conventional beam training [17] has been selected as the baseline method (i.e., communication-only system) to compare proposal one. The algorithm aims to find the optimal receiving beamforming weight by looping the best SNR among all sampling angle directions. Consider a set of M beamforming directions $\boldsymbol{\Theta} = \{\theta_1, \ldots, \theta_M\}$ with corresponding precoders $\mathbf{f}_m = \mathbf{a}_\mathrm{T}^*(\theta_m)$. Considering the channel is static during each time t, the corresponding received vector at the UE can be denoted as:

$$y_{n,m}^{\mathrm{com}} = (\mathbf{h}_n^{\mathrm{com}})^\top \mathbf{f}_m x_{n,m} + z_{n,m}^c. \quad (9)$$

The optimal precoder is selected by maximizing the received power:

$$\hat{m} = \arg\max_m \sum_{n=1}^N |y_{n,m}^{\mathrm{com}}|^2, \quad (10)$$

$$\hat{\theta}^c = \theta_{\hat{m}}, \quad (11)$$

which is then used in (6) for beamforming for communication. This process occupies a certain amount of time $D = M/\Delta f$. Better estimation would be obtained as more time and resources are used to transmit the beams for scanning. However, this process could introduce a high overhead since, in each transmitted frame, the BS sends downlink pilots for beam training and then the data transmission.

B. Proposed Method

According to the localization literature [18], obtaining a coarse estimate of the parameter and using it as a starting point for the process is essential. In this step, monostatic sensing can offer significant assistance without consuming communication resources. Nevertheless, estimating parameters can be complex for scenarios with many obstacles, such as InF, as the echoes from the user need to be more distinguishable from the rest of the echoes in the scenario. To this end, the background subtraction is proposed as a subroutine in the proposed method.
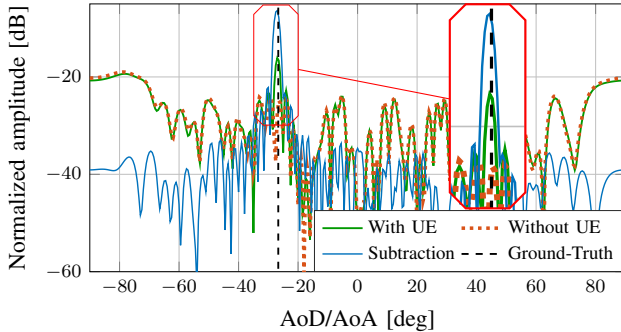

Fig. 2: Beam selection estimation based on background subtraction.

1) Background Subtraction
The estimation of the AoD or AoA based on background subtraction proceeds as follows:

- *Step 1 (Learning stage):* The scenario is previously sensed *without the UE*. The received measurement echo signal at the BS can be expressed as:

$$\mathbf{y}_{n,m}^{\mathrm{ref}} = \tilde{\mathbf{H}}_n \mathbf{f}_m x_{n,m} + \mathbf{z}_{n,m}^s \in \mathbb{C}^{N_\mathrm{R} \times 1}, \quad (12)$$

where $\tilde{\mathbf{H}}_n$ is the sensed channel without considering the UE, i.e., considering echoes only from the rest in the scenario. The pairs $(\mathbf{f}_m, \mathbf{y}_{n,m}^{\mathrm{ref}})$ are stored in a database. The selected beamforming directions cover a fine grid of angles.

- *Step 2 (Inference stage):* Assuming that BS transmits with a precoder $\mathbf{f}$, the received signal can be formulated as in (1).

  - *Step 2a (background subtraction):* Find the index m for which $\mathbf{f}_m$ is as close as possible to the current precoder $\mathbf{f}$, i.e.,

$$\hat{m} = \arg\min_m \|\mathbf{f} - \mathbf{f}_m\|. \quad (13)$$

  From the database, the corresponding $\mathbf{y}_{n,\hat{m}}^{\mathrm{ref}}$ is selected to compute the subtracted signal:

$$\mathbf{y}_n^{\mathrm{sub}} = \mathbf{y}_n^{\mathrm{sen}} - \mathbf{y}_{n,\hat{m}}^{\mathrm{ref}}. \quad (14)$$

  - *Step 2b:* Finally, conventional beamforming is applied to estimate AoD/AoA at monostatic sensing. The best beam to transmit is selected in the angular range $\theta \in [-\pi/2, \pi/2]$. The best angle is obtained by maximizing the subtracted measurement signal $\mathbf{y}_n^{\mathrm{sub}}$, as

$$\hat{\theta}^{\mathrm{sub}} = \arg\max_\theta \sum_{n=1}^N |\mathbf{a}_\mathrm{R}^H(\theta)\mathbf{y}_n^{\mathrm{sub}}|^2, \quad (15)$$

where $\mathbf{a}_\mathrm{R}(\theta)$ is the steering vector defined in (3).

To give a visual evaluation, Fig. 2 shows the background subtraction. The optimal angle is obtained as input for the following time step to construct the predictive beamforming.

2) Predictive Beamforming
Since the UE is moving along a linear trajectory, the Kalman Filter (KF) algorithm can be applied to track the dynamic state information. This enables the precise prediction of the state of the UE, and this information might be utilized in constructing the predictive beamforming. To delve deeper into the process, it is essential to establish an initial state. In this context, the estimated AoD obtained from conventional beam training at $t = 1$ (Section III-A) is considered the initial state as $\hat{\mathbf{x}}_t = (\hat{\theta}^c, v)$, where $\mathbf{x}$ is the state vector filtered in every time step, and v is a preset constant velocity.

The state error covariance matrix should also be determined at the startup. Hence, it can be defined initially as $\mathbf{P} = \mathrm{diag}(\bar{\sigma}_\theta^2, \bar{\sigma}_v^2)$, which would be naturally updated in every tth step. Then, the state vector is predicted using the state evolution model, which in this case is denoted as $\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}\hat{\mathbf{x}}_{t-1}$, where $\mathbf{F}$ is the transition matrix defined by the system dynamics where acceleration has been ignored for simplicity. Next, the process noise covariance matrix is

related to the uncertainty of the predictive results, which can be expressed as $\mathbf{Q} = \mathrm{diag}(\sigma_\theta^2, \sigma_v^2)$.

Following the process of Algorithm 1, another important aspect is the uncertainty information, also called measurement. This information obtains the desired state to update the following time steps. Since the uncertainty can be obtained through an independent system [19], this is where the background subtraction can be used. In this explanation, the estimated AoD from the background subtraction (15) at each time step is detonated as z_t, where the noise covariance R_t is a scalar determined by the variance of the ground truth data.

Algorithm 1: Predictive beamforming

---

1 $t = 1$, perform conventional beam training to obtain $\hat{\theta}^c$.

2 Set the initial state as: $\hat{\mathbf{x}}_t = \begin{bmatrix} \hat{\theta}^c \\ v \end{bmatrix}$

3 for $t \in 2, 3, ...$ **do**

State prediction:

4 $\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}\hat{\mathbf{x}}_{t-1}$

Predicted state covariance matrix:

5 $\mathbf{P}_{t|t-1} = \mathbf{F}\mathbf{P}_{t-1}\mathbf{F}^\top + \mathbf{Q}_{t-1}$

Kalman gain matrix:

6 $\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{h}_t(\mathbf{h}_t^\top \mathbf{P}_{t|t-1}\mathbf{h}_t + R_t)^{-1}$

7   **switch**

8   **case** *Proposed method*

9 $z_t = \hat{\theta}_t^{\mathrm{sub}}$

     end

10   **case** *Data association method*

11 $z_t = \arg\min_j |z_{j,t} - \mathbf{h}_t^\top \hat{\mathbf{x}}_{t|t-1}|$

     end

   end

Innovation (pre-fit residual):

12 $\tilde{y}_t = z_t - \mathbf{h}_t^\top \hat{\mathbf{x}}_{t|t-1}$

Update/correct state:

13 $\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\tilde{y}_t$

Update covariance matrix:

14 $\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{h}_t) + \mathbf{P}_{t|t-1}$

The transmit beamforming vector is constructed by:

15 $\mathbf{f}_t = \mathbf{a}_{\mathrm{T}}^*(\hat{\theta}_t^{\mathrm{pre}}) \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$, as (6) expressed, where $\hat{\theta}_t^{\mathrm{pre}}$ is the predicted angle obtained from $\hat{\mathbf{x}}_t$ (line 13).

end

C. Data Association

For scenarios where background subtraction is not considered, tracking a single UE may be difficult. Given the high clutter density in proposed sub-scenarios, multiple measurements/uncertainties can arise within the area stemming from multiple echoes. In a downlink scenario without uplink feedback from the UE, the BS might need to associate those UE echoes to the predictive state.

To this end, some extensions of linear KF [20] arise in conventional sensing applications. The nearest neighbor method [21] represents one of the simplest approaches, involving the BS in calculating the Euclidean distance between measurements and the predictive state. This method is also outlined in

Scenario layout	
Sub-scenario	Sparse High and Dense High
Room size (WxL)	Small-hall → L = 120 m, W = 60 m
Ceiling height	10 m
BS antenna height	8 m
UE width, length, height	0.2 x 0.2 x 0.2 m
Clutter density	Low clutter density: 20% High clutter density 60%
Clutter height	Low clutter density: 2 m High clutter density: 6 m
Distance between clutter	Low clutter density: 10 m High clutter density: 2 m
UE Trajectory	10 m
Time steps	100
Configuration parameters	
Carrier frequency	28 GHz
Bandwidth	100 MHz
Total transmit power P	21 dBm
Noise variance σ_N^2	10^{-9}
N_T	64
Time frame T	1 ms
Algorithm 1 assumptions	
σ_θ	1 [deg]
σ_v	0.01 [deg/s]
$\bar{\sigma}_\theta$	Based on 3dB beam width
$\bar{\sigma}_v$	0.01 [deg/s]

TABLE I: Indoor factory - simulation assumptions.

[12] as a beam association technique for multiple targets. Due to its significant presence in the literature, this paper selects it as a baseline method within the SAC framework to offer a comprehensive comparison with the proposed approach. The process begins by identifying the most suitable measurement to achieve a reliable predictive state. Therefore, it is assumed that the BS is capable of receiving multiple candidate measurements $z_{j,t}$ from $j = 1, ..., J$ at every t. These J candidate measurements are chosen from the backscattered signal with a constant false alarm rate (CFAR) threshold. In this way, only measurements with similar and highest amplitudes are considered to compute the Euclidean distance between the current a priori prediction and the current observation. In Algorithm 1, line 10 defines the aforementioned process as a *case*, in which the closest measurement state would yield the smallest Euclidean distance, as is described in line 11.

IV. ANALYSIS AND RESULTS

A. Scenario and Channel Model

From the RT tool, accurate channel parameters are derived from detailed geometric data concerning scatterers and their interactions within the environment. The scenario layout follows the recommendation of the 3GPP InF channel model in Release 16 [22]. This version introduces new channel parameters categorized by their industrial geometric conditions. InF channel model incorporates several sub-scenarios classified according to antenna height and clutter density. Specifically, the *Sparse* sub-scenarios signify a low clutter density and are further divided into *Sparse High* and *Sparse Low* relating to high and low BS antenna heights. On the other hand, the *Dense* sub-scenario represents a high clutter density, with its divisions based on antenna height as well, namely *Dense High* and *Dense Low*. For this work, *Sparse high* and *Dense high* sub-scenarios are considered since the UE detection might be
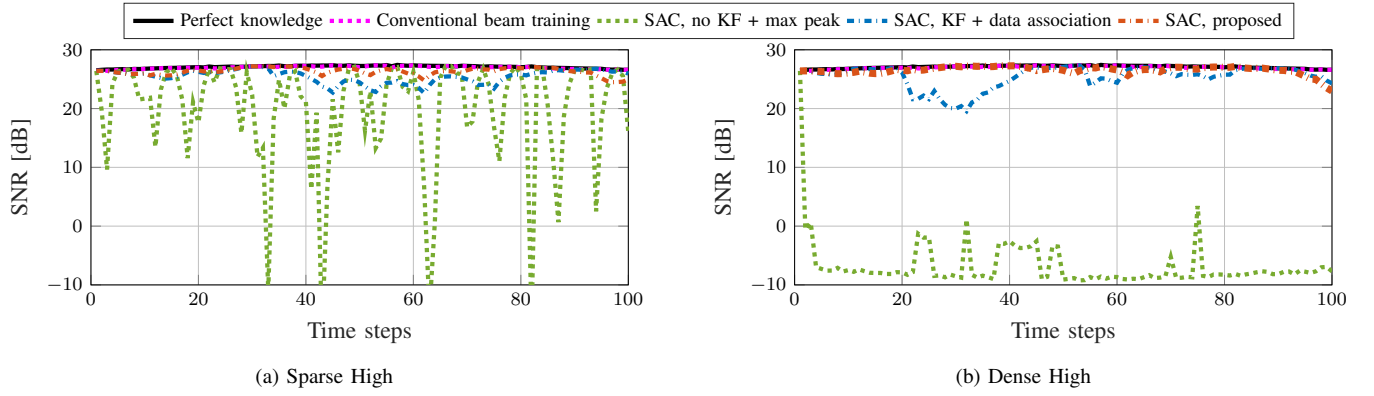
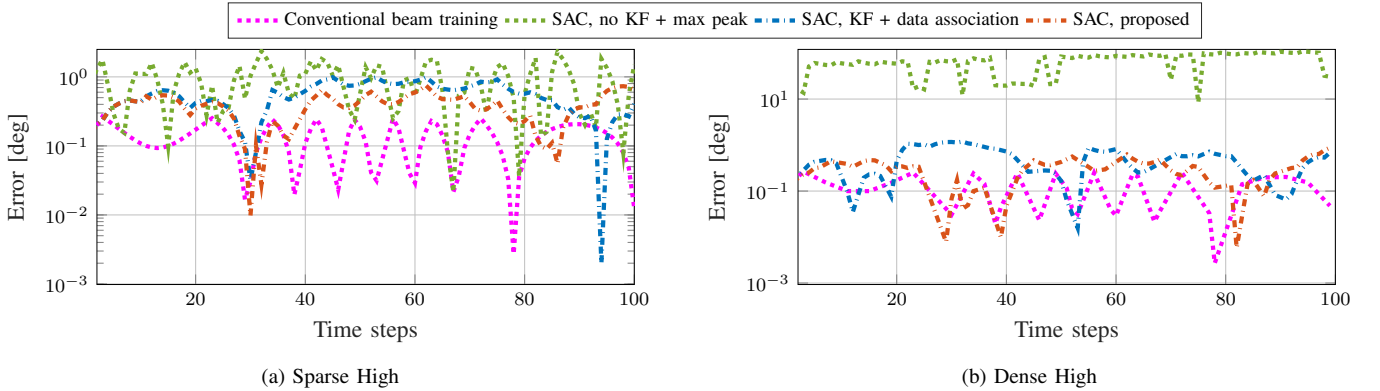Fig. 3: SNR performance for the proposed method and benchmark cases.

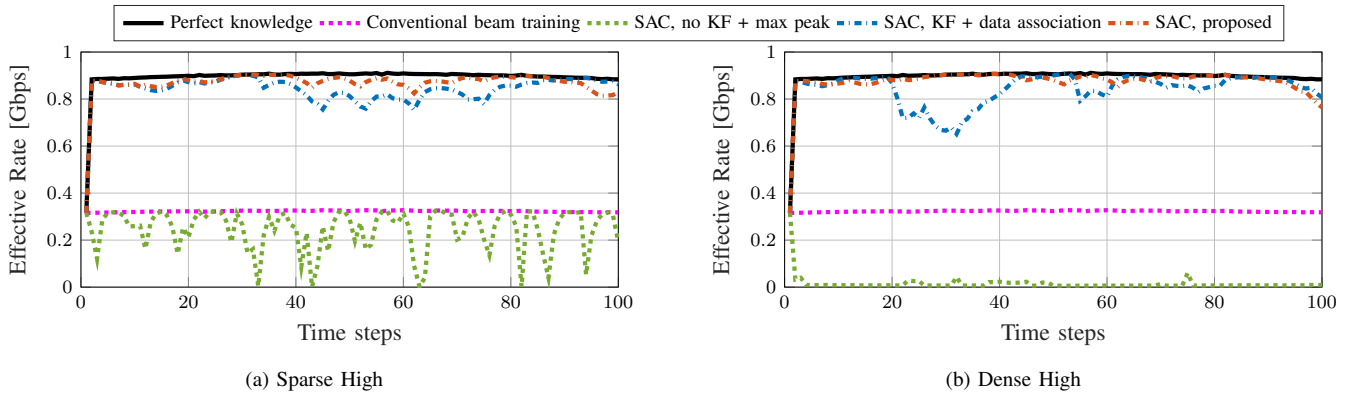Fig. 4: Error performances to give a visual correlation with the low levels of SNR.

Fig. 5: Effective rate performance for the proposed method and benchmark cases.

better performed with the antenna BS higher than the clutter as a first assumption. Both sub-scenarios are constructed by concrete material, where a metal small automated Guided vehicles (AGV) is considered as a UE. The AGV follows a straight-line trajectory of 10 meters in front of the BS in LoS. The acceleration is neglected in these cases since the simulations are snapshots at every time step. From this view, the simulations are performed, and the RT channel dataset is extracted. Table I summarizes the scenario layout and other configuration parameters.

B. Results and Discussion

In order to provide a comparative analysis, several beam training methods are presented in the following results. *Conventional beam training* is the method explained in Sec-

tion III-A. *SAC, proposed* is related to Section III-B, where background subtraction and predictive beamforming is considered. *SAC, KF + data association* is the benchmark of predictive beamforming, which deals with multiple measurements (Section III-C). In addition, *Perfect knowledge* is the ideal case in which the BS can perfectly determine the UE location, discarding the overhead of downlink and uplink pilots in SAC systems. Finally, *SAC, no KF + max peak* determines the optimal beams by finding the angle that maximizes the output of the spatial matched filter applied to the backscattered sensing signal in (1) at every time step without background subtraction (i.e., $\hat{\theta}^{\mathrm{sen}} = \arg\max_{\theta} \sum_{n=1}^{N} |\mathbf{a}_{\mathrm{R}}^{H}(\theta)\mathbf{y}_n^{\mathrm{sen}}|^2$). The latter method has been selected to illustrate the case of misaligned beamforming without filtering since it estimates

the AoD without any predictive method.

Fig. 3 shows the level of SNR over the time steps for all mentioned performances. In both sub-scenarios, the behavior of *Perfect knowledge* and *Conventional beam training* exhibits some similarities. In conventional communication, beam training aims to establish robust links using highly directional beams that precisely align the transmitter beam with the UE. However, this process may introduce high overhead that affects the data transmission. On the other hand, it is anticipated that *SAC, no KF + max peak* will yield low SNR levels. If the optimal beam is selected by the highest power, it might inadvertently capture echo information from a nearby obstacle, such as the ground. This underscores the importance of incorporating techniques such as predictive beamforming and background subtraction to leverage the sensing information. In addition, although the *SAC, KF + data association* obtains a significant level of SNR, the *SAC, proposed* slightly overcomes it. This behavior demonstrates that background subtraction may leverage a strong link in SAC performance.

Key emphasis is that the SNR levels are directly related to detection accuracy. This implies that the accuracy of the information used to direct the transmitted beam should be exceptionally high. Consequently, it is valuable to examine the Error between the ground truth and estimated/predicted information. In Fig. 4, there is a correlation between high levels of error and low levels of SNR in both sub-scenarios. Even though *Dense High* sub-scenario is the most affected, *SAC, proposed* has lower error levels in more time steps than *SAC, KF + data association*.

The overhead is another aspect to analyze. Fig. 5 shows the effective data rate over time steps. Assuming the overhead of conventional beam training should be considered at the first time step, both *SAC, KF + data association* and the *SAC, proposed* obtain high levels of effective data rate. They neglect the overhead for the rest of the time step using predictive beamforming.

V. Conclusion

The evolution of wireless communication systems, transitioning from 5G to the promising realm of 6G, has underscored the significance of precise localization, especially in challenging environments like Indoor factories (InFs). This paper delved into the potential of ISAC systems, emphasizing the coexistence of sensing and communication capabilities. Our findings, derived from realistic measurements using a Ray Tracing tool, highlighted the efficacy of background subtraction in enhancing user detection and predictive beamforming in SAC systems. The results showcased that while conventional beam training methods can introduce significant overhead, SAC, especially with background subtraction, can achieve superior SNR levels and effective data rates.

This paper has the potential to pioneer the use of background subtraction in various estimation algorithms under the ISAC framework. Specifically, it could significantly impact algorithms for detecting distance, velocity, and positioning. Moreover, delving into real-life scenarios that involve multiple users presents a promising avenue for substantial contributions in forthcoming research endeavors.

References

[1] 5G-ACIA, "White Paper: 5G for Connected Industries and Automation," Tech. Rep., February 2019. [Online]. Available: https://5g-acia.org/wp-content/uploads/2021/04/WP_5G_for_Connected_Industries_and_Automation_Download_19.03.19.pdf

[2] T. Jiang *et al.*, "3GPP standardized 5G channel model for IIoT scenarios: A survey," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 8799–8815, 2021.

[3] A. Ramos *et al.*, "Implementation and Calibration of the 3GPP Industrial Channel Model for Ns-3," in *Proceedings of the 2022 Workshop on Ns-3*, ser. WNS3 '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 10–16. [Online]. Available: https://doi.org/10.1145/3532577.3532596

[4] F. Wen *et al.*, "A survey on 5G massive MIMO localization," *Digital Signal Processing*, vol. 94, pp. 21–28, 2019, special Issue on Source Localization in Massive MIMO. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1051200419300569

[5] H. Wymeersch *et al.*, "Integration of Communication and Sensing in 6G: a Joint Industrial and Academic Perspective," in *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2021, pp. 1–7.

[6] F. Lemic *et al.*, "Localization as a feature of mmWave communication," in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2016, pp. 1033–1038.

[7] C. De Lima *et al.*, "Convergent communication, sensing and localization in 6G systems: An overview of technologies, opportunities and challenges," *IEEE Access*, vol. 9, pp. 26 902–26 925, 2021.

[8] 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.

[9] J. A. Zhang *et al.*, "Enabling Joint Communication and Radar Sensing in Mobile Networks—A Survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 306–345, 2022.

[10] H. Wymeersch *et al.*, "Radio Localization and Sensing—Part II: State-of-the-Art and Challenges," *IEEE Communications Letters*, vol. 26, no. 12, pp. 2821–2825, 2022.

[11] Y. Zhang *et al.*, "Toward throughput maximization of integrated sensing and communications enabled predictive beamforming for 6G," *IEEE Network*, pp. 1–1, 2023.

[12] F. Liu *et al.*, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7704–7719, 2020.

[13] Z. Wang *et al.*, "Multi-Vehicle Tracking and ID Association Based on Integrated Sensing and Communication Signaling," *IEEE Wireless Communications Letters*, vol. 11, no. 9, pp. 1960–1964, 2022.

[14] Z. Du *et al.*, "Integrated Sensing and Communications for V2I Networks: Dynamic Predictive Beamforming for Extended Vehicle Targets," *IEEE Transactions on Wireless Communications*, vol. 22, no. 6, pp. 3612–3627, 2023.

[15] Y. Li *et al.*, "ISAC-enabled V2I networks based on 5G NR: How many overheads can be reduced?" *arXiv preprint arXiv:2301.12787*, 2023.

[16] Y. Cui *et al.*, "Seeing Is Not Always Believing: ISAC-Assisted Predictive Beam Tracking in Multipath Channels," *IEEE Wireless Communications Letters*, 2023.

[17] J. Grythe *et al.*, "Beamforming algorithms-beamformers," *Technical Note, Norsonic AS, Norway*, 2015.

[18] A. Shahmansoori *et al.*, "Position and Orientation Estimation Through Millimeter-Wave MIMO in 5G Systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1822–1835, 2018.

[19] M. B. Rhudy *et al.*, "A Kalman filtering tutorial for undergraduate students," *International Journal of Computer Science & Engineering Survey*, vol. 8, no. 1, pp. 1–9, 2017.

[20] B.-n. Vo *et al.*, "Multitarget tracking," *Wiley encyclopedia of electrical and electronics engineering*, no. 2015, 2015.

[21] K. Taunk *et al.*, "A brief review of nearest neighbor algorithm for learning and classification," in *2019 international conference on intelligent computing and control systems (ICCS)*. IEEE, 2019, pp. 1255–1260.

[22] 3GPP, "TR 38.901 V16.0.0: Study on Channel Model for Frequencies from 0.5 to 100 GHz (Release 16)," Tech. Rep., Oct 2019.