# Dynamic Beamforming Design and Antenna Location Optimization for Integrated Sensing and Communication Systems with SWIPT

***Abstract*—**

***Index Terms*—Movable antenna, integrated radar and communication, beamforming, SWIPT, PPO.**

## I. Introduction

## II. System Model

As shown in Fig. , we consider an ISAC with SWIPT which consists of a BS, a set of communication users $\mathcal{K}=\{k|k=1,\dots,K\}$, and a single target. The BS is equipped with $N$ transmitting MA (Tx-MA) elements to serve $K$ fixed single-antenna users, and a single receiving MA (Rx-MA) to perform radar tasks. The positions of Tx-MA and Rx-MA can be flexibly adjusted in given one-dimensional line segments of length $A_t$ and $A_r$, respectively. We define $\mathbf{x}=[x_1,\dots,x_N]^{\mathrm{T}}\in\mathbb{R}^N$ and $x_s$ as the positions of the Tx-MA array and Rx-MA, respectively. In order to prevent coupling of adjacent Tx-MA elements, a minimum distance $D_0$ between antenna pairs is ensured, i.e, $|x_z-x_q|\geq D_0$, $z,q=1,\dots,N,z\neq q$.

We consider that the operation of the ISAC system in multi-time slots. In each time slot, the BS transmits information symbols to the $K$ users along with a radar symbol to the target. We assume that each user is equipped with a simultaneous wireless information and power transfer (SWIPT) unit using power-splitting (PS) protocol. The use of PS-based SWIPT protocol allows the user to split the signal received from the BS into two parts. The first part is for the energy harvesting, which helps to prolong its operation lifetime, and the second part is for the information decoding. We denote $\rho$ as is the ratio of the received signal power used to harvest energy. Therefore, $(1-\rho)$ is the received signal power used to decode the information.

### A. Channel Model

In this paper, we assume that the far-field condition is satisfied since the region for antenna movement is much smaller than the signal propagation distance. Therefore, all the Tx-MA elements experience the same angle of direction (AoD) $\varphi\in[0,\pi]$. We consider that there are $C$ clusters in the system. We denote $\mathbf{a}_s(\mathbf{x})$, $\mathbf{a}_c(\mathbf{x})$, $\mathbf{a}_{k,l}(\mathbf{x})\in\mathbb{C}^N$ as the field-response vectors (FRVs) between the BS and the target/clutter $c$, and the $l$-th propagation path from the BS to user $k$, respectively, which can be written as

$$\mathbf{a}_s(\mathbf{x})=[e^{j\frac{2\pi}{\lambda}x_1\cos\varphi_s},\dots,e^{j\frac{2\pi}{\lambda}x_N\cos\varphi_s}]^{\mathrm{T}},\tag{1}$$

$$\mathbf{a}_c(\mathbf{x})=[e^{j\frac{2\pi}{\lambda}x_1\cos\varphi_c},\dots,e^{j\frac{2\pi}{\lambda}x_N\cos\varphi_c}]^{\mathrm{T}},\tag{2}$$

$$\mathbf{a}_{k,l}(\mathbf{x})=[e^{j\frac{2\pi}{\lambda}x_1\cos\varphi_{k,l}},\dots,e^{j\frac{2\pi}{\lambda}x_N\cos\varphi_{k,l}}]^{\mathrm{T}},\tag{3}$$

where $\varphi_s$, $\varphi_c$ and $\varphi_{k,l}$ denote the AoD of target/clutter $c$ and $l$-th path of user $k$, respectively, and $\lambda$ is the wavelength.

We assume that perfect channel state information (CSI) is available at the BS [1]. Let $L_p$ denote the total number of channel paths from the BS to user $k$. Then, the channel vector between the BS and user $k$ is obtained as

$$\mathbf{h}_k=\sum\nolimits_{l=1}^{L_p}\beta_{k,l}\mathbf{a}_{k,l}(\mathbf{x})\in\mathbb{C}^{N\times1},\tag{4}$$

where $\beta_{k,l}\in\mathbb{C}$ denotes the path gain of the l-th transmit path from the BS to user $k$. We assume that $\beta_{k,l}$ follows the circularly symmetric complex Gaussian distribution [2], i.e., $\beta_{k,l}\sim\mathcal{CN}(0,\delta d_k^{-\alpha}/L_p)$, where $\delta$ represents the path loss at the reference distance of 1 m and $\alpha$ is the path-loss exponent. The BS simultaneously transmits information symbols to $K$ users and dedicated radar signals to the target. We define $\mathbf{s}=[s_1,\dots,s_K,s_{K+1}]^{\mathrm{T}}\in\mathbb{C}^{K+1}$ as the transmit signal vector from the BS, where $s_k$, $k\in\mathcal{K}$, is the information symbol attended to the users $k$, respectively, and $s_{K+1}$ is the dedicated radar signal for sensing, and $\mathbb{E}[\mathbf{s}^{\mathrm{H}}\mathbf{s}]=\mathbf{I}$. The BS employs the linear beamformer to construct the signal from BS. We denote $\mathbf{W}=[\mathbf{w}_1,\dots,\mathbf{w}_K,\mathbf{w}_{K+1}]\in\mathbb{C}^{N\times(K+1)}$ as the beamforming matrix of BS. Then, the received signal at the user $k$ is given by

$$y_k=\underbrace{\mathbf{h}_k^H(\mathbf{x})\mathbf{w}_k s_k}_{\text{expected signal}}+\underbrace{\sum\nolimits_{j=1,j\neq k}^{K}\mathbf{h}_k^H(\mathbf{x})\mathbf{w}_j s_j}_{\text{inter-user interference}}+\underbrace{\mathbf{h}_k^H(\mathbf{x})\mathbf{w}_{K+1}s_{K+1}}_{\text{sensing interference}}+n_k\tag{5}$$

where $n_k\sim\mathcal{CN}(0,\sigma_k^2)$ is the additive white Gaussian noise (AWGN) with mean zero and variance $\sigma_k^2$. The data rate at user $k$ is given by

$$R_k(\mathbf{W},\mathbf{x},\rho)=\log_2(1+\frac{(1-\rho)\|\mathbf{h}_k^H(\mathbf{x})\mathbf{w}_k\|^2}{\psi_k(\mathbf{W},\mathbf{x},\rho)}).\tag{6}$$

where $\psi_k(\mathbf{W},\mathbf{x},\rho)=(1-\rho)\sum_{j=1,j\neq k}^{K+1}\|\mathbf{h}_k^H(\mathbf{x})\mathbf{w}_j\|^2+\sigma_k^2$.

The BS transmits radar signal and then receives echo signal from target and clutters. The radar operates in track model, where an initial estimate of the target's AoA is available from previous scanning. We model the channel between the BS and target denoted by $\mathbf{g}_s\in\mathbb{C}^N$, varying slowly with time, with its expression [3]:

$$\mathbf{g}_s=\beta_s e^{j2\pi f_d\tau}e^{j\frac{2\pi}{\lambda}x_s\cos\varphi_s}\mathbf{a}_s(\mathbf{x}),\tag{7}$$

where $f_d=2vf_c/v$ is the Doppler frequency, $f_c$ and $v$ are the signal carrier frequency and the speed of light, respectively,

$\tau$ represents the system sampling period, and $\beta_s$ is the attenuation coefficient, including round-trip path-loss and radar cross-section (RCS) $\alpha_s$, which given by

$$\beta_s = \sqrt{\frac{\lambda^2 \alpha_s}{(4\pi)^3 d_s^4}}. \tag{8}$$

Additionally, $\mathbf{g}_c \in \mathbb{C}^N$ is the channel between the BS and clutters $c$ can be derived similar to (7). Then, the echo signal for sensing can be written as

$$y_s = \underbrace{\mathbf{g}_s^H \mathbf{w}_{K+1} \mathbf{s}_{K+1}}_{\text{target}} + \underbrace{\sum_{j=1}^{K} \mathbf{g}_s^H \mathbf{w}_j \mathbf{s}_j}_{\text{users}} + \underbrace{\sum_{c=1}^{C} \sum_{i=1}^{K+1} \mathbf{g}_c^H \mathbf{w}_i \mathbf{s}_i}_{\text{clutters}} + n_s, \tag{9}$$

where $n_s \sim \mathcal{CN}(0, \sigma_s^2)$ is the AWGN with mean zero and variance $\sigma_s^2$. Thus, we can obtain the radar signal-to-clutter-plus-noise ratio (SCNR) at the BS as

$$\text{SCNR}(\mathbf{W}, \mathbf{x}, x_s) = \frac{\|\mathbf{g}_s^H(\mathbf{x}) \mathbf{w}_{K+1}\|^2}{\psi_s(\mathbf{W}, \mathbf{x}, x_s, \rho)}. \tag{10}$$

where $\psi_s(\mathbf{W}, \mathbf{x}, x_s) = \sum_{j=1}^{K} \|\mathbf{g}_s^H(\mathbf{x}) \mathbf{w}_j\|^2 + \sum_{c=1}^{C} \sum_{i=1}^{K+1} \|\mathbf{g}_c^H(\mathbf{x}) \mathbf{w}_i\|^2 + \sigma_s^2$.

### B. Wireless Energy Harvesting

The users harvest part of signal energy received from BS, which helps to prolong its operation lifetime. In this paper, we consider a realistic scenario in which the non-linear energy harvesting model [4] is used, i.e., at the users. Thus, the amount of harvested energy of user $k$ is $\rho E_k(\mathbf{W}, \mathbf{x}, \rho)$, where

$$E_k(\mathbf{W}, \mathbf{x}, \rho) = \frac{\Gamma_k - \vartheta_k \Upsilon_k}{1 - \Upsilon_k}, \tag{11a}$$

$$\Upsilon_k = \frac{1}{1 + \exp(\varrho_k \gamma_k)}, \tag{11b}$$

$$\Gamma_k = \frac{\vartheta_k}{1 + \exp(-\varrho_k (P_k^{\text{EU}} - \gamma_k))}, \tag{11c}$$

where $P_k^{\text{EU}} = \sum_{k=1}^{K+1} |\mathbf{h}_k^{\text{H}}(\mathbf{x}) \mathbf{w}_k|^2$ is the received power at user $k$, $\Gamma_k$ is the logistic function for a given received power at user $k$ and $\vartheta_k, \gamma_k, \varrho_k$ are constant parameters determined by the employed energy harvesting circuit [5]. Here, we ignore the energy harvested from the noise.

### C. Problem Formulation

The goal of the system is to maximize sum transmition data in each time step. Mathematically, the flexible beamforming optimization problem is formulated as:

$$\max_{\rho, \mathbf{x}, \mathbf{W}} \quad \sum_{k=1}^{K} R_k \tag{12a}$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{W}^H \mathbf{W}) \leq P_0 \tag{12b}$$

$$0 \leq x_n \leq A_t, \quad 1 \leq n \leq N, \tag{12c}$$

$$0 \leq x_s \leq A_r, \tag{12d}$$

$$|x_z - x_q| \geq D_0, \quad z, q = 1, \ldots, N, z \neq q, \tag{12e}$$

$$\text{SCNR} \geq \text{SCNR}_{\min}, \tag{12f}$$

$$\rho E_k \geq E_{\min}, \quad \forall k \in \mathcal{K}, \tag{12g}$$

$$\rho \in [0, 1]. \tag{12h}$$

## III. Stochastic Optimization Problem Formulation

The optimization problem formulated in Section II is reformulated as a Markov decision process (MDP). This representation is defined by the tuple $\langle S, A, P, R \rangle$, where $S$, $A$, $P$, and $R$ denote the state space, action space, state transition probability function, and reward function, respectively.

### A. State Space

The state space, $S$, represents the environment's status at each time step. It includes:

- Antenna positions $\mathbf{x}$ for both Tx-MA and Rx-MA arrays.
- Received signal power at each user $P_k$.
- Current beamforming weights $\mathbf{W}$.
- Radar signal-to-clutter-plus-noise ratio (SCNR).

### B. Action Space

The action space, $A$, consists of possible adjustments to:

- Antenna positions within the movement constraints, $x_z, x_q \in \mathcal{X}$, ensuring $\|x_z - x_q\| \geq D_0$.
- Beamforming matrix $\mathbf{W}$ for optimizing signal power distribution.
- Power-splitting ratio $\rho$ for SWIPT users.

### C. State Transition Function

The state transition probability function, $P$, models the dynamics of the environment in response to the selected actions. For instance:

$$P(s_{t+1} | s_t, a_t) = f(s_t, a_t),$$

where $f$ encapsulates the physics of antenna movement, signal propagation, and energy harvesting.

### D. Reward Function

The reward function, $R$, quantifies the performance of the system in achieving the objectives. It balances the competing goals of:

- Maximizing the sum data rate for all users, $\sum_{k=1}^{K} R_k(W, \mathbf{x}, \rho)$.
- Ensuring SCNR requirements for radar operations: $\text{SCNR} \geq \text{SCNR}_{\min}$.

- Meeting the harvested energy threshold for all users: $\rho E_k(W, \mathbf{x}, \rho) \geq E_{\min}$.

The cumulative reward over a time horizon $T$ is given by:

$$R_{\text{total}} = \sum_{t=1}^{T} \gamma^{t-1} R_t,$$

where $\gamma \in [0, 1]$ is the discount factor for future rewards.

### *E. Reformulated Problem*

The MDP-based reformulation of the original problem aims to find an optimal policy $\pi^*$:

$$\pi^* = \arg\max_{\pi} \mathbb{E}\left[R_{\text{total}} \mid \pi\right],$$

where $\pi$ represents a mapping from states to actions, $\pi : S \rightarrow A$.

This reformulation enables the use of advanced reinforcement learning techniques, such as Proximal Policy Optimization (PPO), to derive efficient solutions.

## IV. DRL Algorithm

## V. Performance Evaluation

SCNR and energy harvested thresholds. $E_{\min} = \{-15, -20, -25, -30\}$ dBm, $SCNR_{\min} = \{5, 10, 15\}$.

Power $P_0 = \{20, 25, 30, 35\} dbm$.

The number of users $K = \{2, 3, 4, 5\}$.

algorithm parameter change comparison table

The users and clutters are randomly distributed within the range of $[0, \pi]$, and the target is located at $60^0$.

TABLE I: Simulation Parameter

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| $K$ | 3 | $C$ | 2 |
| $N$ | 4 | $\varphi_s$ | $60^0$ |
| $\tau$ | 1 $\mu s$ [6] | ? | ? |
| $L_p$ | 13 | $\lambda$ | 0.1m |
| $\alpha$ | 2.2 | $D_0$ | $\lambda/2$ |
| $\sigma_k^2, \sigma_s^2$ | $-80$ dBm | $\vartheta_k$ | 20 [5] |
| $\gamma_k$ | 8 [5] | $\varrho_k$ | 0.3 [5] |
| $E_{\min}$ | $-20$ dBm ??? | $P_o$ | 32 dBm |
| $A$ | $10 \times \lambda$ | $\text{SCNR}_{\min}$ | $5 - 10$ dBm??? |
| $\delta_k$ | 1 | $\nu$ | 3.6 |
| $d_{\max}$ | 100 | $d_{\min}$ | 25 |
| $f_c$ | 2.4 GHz | $v$ | $3 \times 10^8$m/s |
| $\delta$ | 1 | $\alpha$ | 2.8 |

## VI. Conclusions

## References


[1] L. Zhu, W. Ma, B. Ning, and R. Zhang, "Movable-antenna enhanced multiuser communication via antenna position optimization," 2024. [Online]. Available: https://arxiv.org/abs/2302.06978

[2] X. Wei, W. Mei, D. Wang, B. Ning, and Z. Chen, "Joint Beamforming and Antenna Position Optimization for Movable Antenna-Assisted Spectrum Sharing," *IEEE Wireless Communications Letters*, vol. 13, no. 9, p. 2502–2506, 2024.

[3] Z. Liu, S. Aditya, H. Li, and B. Clerckx, "Joint Transmit and Receive Beamforming Design in Full-Duplex Integrated Sensing and Communications," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 9, pp. 2907–2919, 2023.

[4] D. Xu, X. Yu, V. Jamali, D. W. K. Ng, and R. Schober, "Resource allocation for large irs-assisted swipt systems with non-linear energy harvesting model," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, 2021, pp. 1–7.

[5] E. Boshkovska, "Practical non-linear energy harvesting model and resource allocation in SWIPT systems," *CoRR*, vol. abs/1602.00833, 2016. [Online]. Available: http://arxiv.org/abs/1602.00833

[6] S. Huang, M. Zhang, Y. Gao, and Z. Feng, "Mimo radar aided mmwave time-varying channel estimation in mu-mimo v2x communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 11, pp. 7581–7594, 2021.