Joint Beamforming and Trajectory Optimization for UAV-Aided ISAC with Dipole Antenna Array

Mustafa Burak Yilmaz, Lin Xiang, and Anja Klein
Communications Engineering Lab, Technical University of Darmstadt, Darmstadt, Germany
Email: {b.yilmaz, l.xiang, a.klein}@nt.tu-darmstadt.de

***Abstract*—This paper explores energy-aware joint optimization of beamforming and trajectory for integrated sensing and communication (ISAC) using an energy-limited unmanned aerial vehicle (UAV). Equipped with a uniform linear array of half-wavelength dipole antennas, the UAV transmits information-carrying signals to simultaneously serve downlink communication users and sense ground targets during its mission. Our aim is to maximize the accumulated sensing energy for the ground targets without violating the energy budget, while ensuring quality-of-service for the communication users by jointly optimizing the UAV's flight trajectory, ISAC beamforming, and mission completion time. The problem we address is inherently nonconvex and typically challenging to solve to optimality. Drawing inspiration from approximate dynamic programming (DP) methods, we propose a novel, computationally efficient solution by combining the one-step lookahead rollout algorithm from approximate DP with semidefinite programming techniques from convex optimization. Simulation results demonstrate that, when compared to two baseline schemes, our proposed approach significantly expands the achievable performance region for both sensing and communication.**

I. Introduction

Integrated sensing and communication (ISAC) is an emerging technology to enable simultaneous sensing and communication in the sixth-generation (6G) wireless networks. By transmitting common signals over shared spectrum and transmitter hardware for both sensing and communication purposes [1], ISAC facilitates highly efficient utilization of the available spectrum, hardware, and processing resources. So far, a substantial body of existing literature has considered the design and optimization of ISAC, such as transmit beamforming optimization [2], [3], primarily in two-dimensional terrestrial wireless networks. It was until recently that extending the ISAC functionality to three-dimensional non-terrestrial networks (NTNs) has gained paramount importance. The latter is motivated by the increasing demand for e.g. providing temporary or emergency communication and sensing service for users outside the coverage of ground networks or in the aftermath of ground disasters [4].

Motivated by this trend, in this paper, we consider ISAC enabled by unmanned aerial vehicles (UAVs). UAVs have inspired versatile onboard sensing and communication applications, to enhance or supplement terrestrial and/or space-based services [5]. Meanwhile, compared with ground or satellite networks, UAVs can be swiftly deployed on demand and easily relocated at relatively lower costs for improved sensing accuracy and communication coverage [5]. Therefore, UAV-aided ISAC is a promising solution for facilitating advanced sensing and communication applications, including UAV-aided simultaneous collection and transmission of fresh sensed data [6], joint localization and communication [7], and tracking-assisted secure communication [8], in 6G NTNs.

As UAVs typically have a limited battery, *how to maximize the sensing and communication performances within the UAV's finite energy supply* defines a key research problem for UAV-aided ISAC. Yet, concrete problem formulations following this aim are often nonconvex, which hinders an efficient solution. In [9], the authors developed an iterative method based on successive convex approximation (SCA) to optimize UAV's flight trajectory and transmit beamforming for maximizing the communication data rate while achieving required beam pattern gains on sensing targets. It was shown in [9] that the proposed method can exploit UAV mobility to simultaneously improve the sensing and communication performances. The authors of [10] employed exhaustive search to optimize the selection of beamforming vectors from a given codebook for maximizing the communication data rate based on sensed information about the propagation environment. However, these works [9], [10] ignored the UAV's energy consumed for propulsion flight. In [11], the authors jointly optimized transmit beamforming, UAV trajectory planning, and sensing schedule to minimize the overall energy consumption of the UAV within a specified service duration, whereas sensing activities were limited to predetermined fixed locations. The authors solved the formulated nonconvex mixed-integer nonlinear programming problem using successive convex approximation (SCA).

For problem tractability, the aforementioned studies [6]–[11] have assumed ideal isotropic antennas or arrays with isotropic antennas for UAV-aided ISAC scenarios. In this paper, we aim to extend the investigation to UAV-aided ISAC utilizing practical directional antennas. In particular, we equip the UAV with a uniform linear array (ULA) of half-wavelength dipole antennas for sending ISAC signals. By harnessing both the gain pattern of directional antennas and the array beamforming, narrowly focused beams can be created to enhance communication and sensing performance while minimizing interference and clutters emitted and intercepted from side-lobes.

Meanwhile, exploiting the UAV's 3D mobility, we further jointly optimize beamforming using the dipole array and trajectory design for the UAV's flight to maximize the sensing performance while satisfying both the quality-of-service (QoS)

This work has been funded by the LOEWE initiative (Hessen, Germany) within the emergenCITY center and has been supported by the BMBF project Open6GHub under grant 16KISKO14.

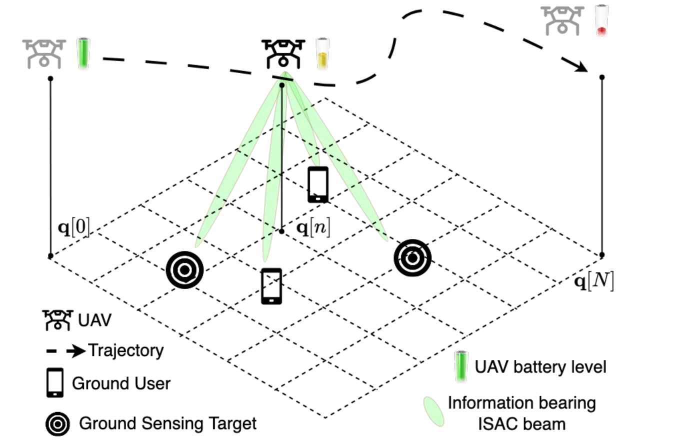

Fig. 1: System model of UAV-aided ISAC with limited onboard battery.

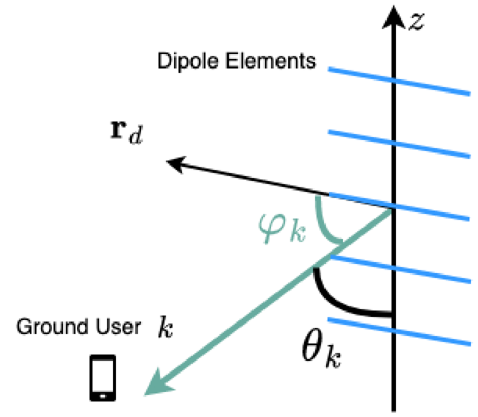

Fig. 2: Departure and azimuth angles of GU k relative to the UAV.

requirements of communication users and a total energy budget for the UAV. Due to the nonlinear gain pattern of dipoles, the resulting optimization problem is more difficult to solve than the ones in [6]–[11]. To tackle this challenge, we propose a novel computationally efficient solution based on the one-step lookahead rollout (OSLR) algorithm from approximate dynamic programming (DP). Our contributions are:

- We investigate joint trajectory design and transmit beamforming optimization for UAV-aided ISAC employing a dipole antenna array, by considering limited onboard energy and limited yet unknown mission time for the UAV. The aim is to maximize the accumulated sensing performance over the whole mission while satisfying communication, flight and energy requirements.
- The problem that has been defined is exceedingly non-convex. Considering the achievements of approximate dynamic programming (DP), we introduce an innovative and efficient approach that combines the one-step lookahead rollout (OSLR) technique from approximate DP with semidefinite programming (SDP).
- Simulation results show that the proposed OSLR algorithm can jointly exploit the UAV mobility and the radiation pattern of dipole antennas for beamforming of the ISAC signals. Consequently, the proposed scheme can significantly improve both sensing and communication performances, outperforming two baseline schemes.

In the remainder of this paper, we present the system model in Section II. The joint optimization problem of trajectory design and dipole array beamforming in UAV-aided ISAC is formulated and solved via the proposed algorithm in Sections III and IV, respectively. Section V provides the simulation results and finally, Section VI concludes the paper.

Notations: Throughout this paper, matrices and vectors are denoted by boldface capital and lower-case letters, respectively. $\mathbf{A}^T$, $\mathbf{A}^H$, Rank($\mathbf{A}$), and Tr($\mathbf{A}$) denote the transpose, Hermitian conjugate transpose, rank, and trace of matrix $\mathbf{A}$, respectively. Finally, $|\cdot|$ and $||\cdot||$ denote the absolute value of a complex scalar and the Euclidean norm of a complex vector, respectively.

II. System Model

In this section, we present the system model for UAV-aided ISAC system using a dipole antenna array. We also introduce the channel and signal models, and the energy consumption model of the studied system.

A. UAV-aided ISAC

Fig. 1 illustrates the UAV-aided ISAC system considered in this paper. In particular, a rotary-wing UAV is deployed to simultaneously communicate with K ground users (GUs) and sense M ground targets (GTs). The UAV performs mono-static sensing with full duplex capability. The locations of the GUs and the GTs are fixed and given by $\mathbf{u}_k^{\mathrm{c}} \in \mathbb{R}^{2\times 1}$, $k = 1, ..., K$, and $\mathbf{u}_m^{\mathrm{s}} \in \mathbb{R}^{2\times 1}$, $m = 1, ..., M$, respectively. Due to a limited onboard battery, the UAV should complete the whole mission within a finite energy budget E^{tot}. Without loss of generality, we assume that the UAV flies at a fixed altitude H during its mission. To facilitate trajectory planning, we consider a continuous-time system, and quantize the target area into a uniform grid. The uniform grid consists of G grid nodes, defined by set $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_G\} \in \mathbb{R}^{2\times 1}$, which constitute the candidate waypoints for UAV's flight. Then, the UAV's flight path is composed of N line-segments defined by $N+1$ waypoints $\mathbf{q}[n] \in \mathcal{V}$, cf. Figure 1, where

$$||\mathbf{q}[n+1] - \mathbf{q}[n]|| \leq \Delta_{\max}, \quad n = 0, \ldots, N. \tag{1}$$

We require $\Delta_{\max} \ll H$ such that the distances between the UAV and each GU/GT barely change when the UAV flies between two waypoints. Hence, we assume that the UAV flies at a constant speed $V[n]$ on each line-segment n, .i.e. the flight segment between waypoints $\mathbf{q}[n]$ and $\mathbf{q}[n-1]$. The UAV flight speed $V[n]$ on line-segment n is defined as $V[n] = ||\mathbf{q}[n] - \mathbf{q}[n-1]||/\tau[n]$ and $\tau[n]$ is the flight time over line-segment n. Note that both G and N are specified a priori, whereas $\tau[n]$ is an adjustable variable. The total length of the line-segments, $N\Delta_{\max}$ gives a upper bound on the total flight distance. Hence this bound is tightened by selecting a sufficiently large N value, but for computational savings an extremely large N is avoided.

B. 3D Channel Model

The dual-functional UAV is equipped with a transmit ULA consisting of $S \geq 1$ half-wave dipole antennas, and each GU employs a single receive antenna. We assume that a strong line-of-sight (LoS) link exists between the elevated UAV and each GU. The channel vector from the UAV to GU k on line-segment n, denoted by $\mathbf{h}_k[n] \in \mathbb{C}^{S\times 1}$, is modeled as

$$\mathbf{h}_k[n] = \frac{\sqrt{\beta}}{\sqrt{\|\mathbf{d}_k[n]\|^2 + H^2}} \cdot \mathbf{a}_k[n], \tag{2}$$

where β denotes the channel power gain at a unit distance and $\mathbf{d}_k[n] \triangleq (\mathbf{u}_k^{\mathrm{c}} - \mathbf{q}[n]) \in \mathbb{R}^{2\times 1}$ is the horizontal displacement vector between the UAV and GU k on line-segment n.

In (2), $\mathbf{a}_k[n]$ is the steering vector of the ULA at the UAV on line-segment n. For convenience of modeling $\mathbf{a}_k[n]$, we assume that the ULA is vertically placed on the UAV as in [9], [11]. A vertical ULA enables beamforming based on the distances from the UAV to the GUs/GTs, and provides more opportunities for transmission to exploit through the design of the trajectory. Moreover, each dipole points toward a direction specified by unit vector $\mathbf{r}_d \in \mathbb{R}^{2\times 1}$, cf. Fig. 2, where $\mathbf{r}_d$ is assumed to be fixed during the UAV's flight. Consequently, while the UAV flies over line-segment n, the angle of departure $\theta_k[n]$, and azimuth angle $\varphi_k[n]$ corresponding to GU k are given as

$$\theta_k[n] = \arctan\left(\frac{\|\mathbf{d}_k[n]\|}{H}\right), \text{ and} \tag{3}$$

$$\varphi_k[n] = \arccos\left(\mathbf{d}_k[n]^T \mathbf{r}_d\right), \tag{4}$$

respectively. Finally, considering the radiation pattern of the dipoles, the steering vector toward GU k is given as

$$\mathbf{a}_k[n] \triangleq f_e(\varphi_k[n]) \cdot [1, e^{j\xi_k[n]}, \ldots, e^{j\xi_k[n](S-1)}]^T, \tag{5}$$

where we define $\xi_k[n] \triangleq 2\pi \frac{d^{\mathrm{ULA}}}{\lambda} \cos(\theta_k[n])$ with d^{ULA} being the spacing between adjacent dipoles and λ the carrier wavelength. Moreover, $f_e(\varphi_k[n])$ denotes the directional dependence, cf. (4), of the radiated electric fields in each dipole, given as

$$f_e(\varphi_k[n]) = \frac{\cos(\frac{\pi}{2}\cos(\varphi_k[n]))}{\sin \varphi_k[n]} \alpha. \tag{6}$$

Here, the normalization coefficient α ensures that the total power radiated from the dipole antenna array in all azimuth angles is equal to power radiated from a ULA of isotropic elements, i.e. $\frac{1}{2\pi}\int_{-\pi}^{\pi} f_e^2(\varphi_k) d\varphi_k = 1$. Note that, as $\Delta_{\max} \ll H$, we assume that the distances, the steering vectors, and the channels between the UAV and each GU/GT all change barely when the UAV flies along each line segment.

C. Performance Metrics of Sensing and Communication

We assume that the information bearing symbol $c_i \in \mathbb{C}$ intended for GU i is a complex Gaussian random variable with zero mean and unit variance. Moreover, let $\mathbf{w}_i \in \mathbb{C}^{S\times 1}$ be the beamforming vector of the UAV for sending c_i. The signal transmitted by the UAV on line-segment n is given as $\mathbf{x}[n] = \sum_{i=1}^{K} \mathbf{w}_i[n]\, c_i[n]$. Subsequently, the received signal at GU k on line-segment n, denoted by $\mathbf{y}_k[n]$, is given as

$$\mathbf{y}_k[n] = \mathbf{h}_k^H[n]\,\mathbf{x}[n] + \eta_k[n], \tag{7}$$

where $\eta_k[n]$ is the additive white Gaussian noise at GU k and is modeled as a zero-mean Gaussian random variable with variance σ_k^2. Based on (7), the achievable data rate of GU k on line-segment n is given as

$$R_k[n] = B\log_2\left(1 + \frac{|\mathbf{h}_k^H[n]\mathbf{w}_k[n]|^2}{\sum_{i=1, i\neq k}^{K} |\mathbf{h}_k^H[n]\mathbf{w}_i[n]|^2 + \sigma_k^2}\right), \tag{8}$$

where B is the system bandwidth in Hz.

We investigate radar sensing towards GTs by utilizing the communication signal $\mathbf{x}[n]$ transmitted via ULA as probing signals to maximize the energy at the locations of GTs. The UAV collects echoes of this signal, reflected or scattered by the GTs, for tasks such as user detection, area monitoring, and GT distinction based on e.g. angle of arrival and round-trip times of the echoes, as detailed in [9] and [1]. Our goal is to improve the accumulated sensing energy for the GTs. Similar to communication, we assume a LoS dominant link with path loss model as given in (2). Consequently, we model the accumulated sensing energy for GT m on line-segment n as

$$\Gamma_m[n] = \frac{\beta P_m^{\mathrm{s}}[n]}{\|\mathbf{d}_m[n]\|^2 + H^2}\, \tau[n], \tag{9}$$

where $P_m^{\mathrm{s}}[n]$ is the power of the transmitted signal in the direction of $\mathbf{u}_m^{\mathrm{s}}$ on line-segment n, given by

$$P_m^{\mathrm{s}}[n] = \sum\nolimits_{k=1}^{K} |\mathbf{w}_k^H[n]\mathbf{a}_m[n]|^2. \tag{10}$$

D. Energy Consumption of the UAV

During the mission, the UAV consumes energy in both signal transmission and propulsion flight. The UAV's energy consumption for signal transmission on line-segment n, denoted by $E^{\mathrm{c}}(\mathbf{w}_k[n])$, is given as

$$E^{\mathrm{c}}(\mathbf{w}_k[n]) = \sum\nolimits_{k=1}^{K} \left(\mathbf{w}_k^H[n]\mathbf{w}_k[n] + P_{\mathrm{const}}\right) \tau[n], \tag{11}$$

which includes the transmit power in the first term and a constant power, P_{const}, consumed in the circuitry and signal processing etc. Meanwhile, the propulsion energy consumption of the rotary-wing UAV on line-segment n is modeled as [12]

$$E^{\mathrm{f}}(V[n]) = \left[P_0\left(1 + \frac{3V[n]^2}{U_{\mathrm{tip}}^2}\right) + P_i\left(\sqrt{1 + \frac{V[n]^4}{4V_0^4}} - \frac{V[n]^2}{2V_0^2}\right)^{1/2} + \frac{1}{2} d_0 \rho s A V[n]^3\right] \tau[n] \tag{12}$$

with rotor disc area A, tip speed of the rotor blade U_{tip}, rotor solidity s, air density ρ, fuselage drag ratio d_0, mean rotor velocity induced in forward flight V_0, blade profile power during hovering P_0, and induced power during hovering P_i, cf. [12].

III. Problem Formulation

For the considered UAV-aided ISAC system, the spatial degrees-of-freedom (DoFs) of the transmit dipole array and the mobility of the UAV can be jointly exploited to maximize the system performance within the given energy budget. In order to properly illuminate all the GTs, we use the worst-case accumulated sensing energy among all GTs, cf. (9), as the sensing performance metric [10]. Assuming that GU and GT locations are known, we formulate the energy-aware trajectory and beamforming optimization problem for UAV-aided ISAC as

$$\text{P1}: \max_{\mathbf{q}[n],\mathbf{w}_k[n],\tau[n]} \ \min_{m} \ \sum\nolimits_{n=1}^{N} \Gamma_m[n] \tag{13}$$
$$\begin{aligned}
\text{s.t. } & \text{C1: } R_k[n] \ge R_{\min,k} \ \ \forall k, \forall n \\
& \text{C2: } \sum\nolimits_{n=1}^{N} \left(E^{\text{c}}(\mathbf{w}_k[n]) + E^{\text{f}}(V[n]) \right) \le E^{\text{tot}} \\
& \text{C3: } \sum\nolimits_{k=1}^{K} \mathbf{w}_k^H[n]\mathbf{w}_k[n] \le P_{\max}, \forall n \\
& \text{C4: } V[n] \le V_{\max}, \forall n \\
& \text{C5: } \mathbf{q}[0] = \mathbf{q}_I, \quad \mathbf{q}[N] = \mathbf{q}_F.
\end{aligned}$$

In problem P1, constraint C1 ensures that, during the ISAC mission, the UAV can continuously communicate with GU k at a data rate being equal to or above $R_{\min,k}$, to guarantee communication QoS. C2 defines the total energy budget for the UAV mission. C3 and C4 limit the maximum transmit power and the maximum flight speed of the UAV by $P_{\max}$ and $V_{\max}$, respectively. Finally, C5 specifies the starting and destination positions for the UAV flight to be $\mathbf{q}_I \in \mathbb{R}^{2\times 1}$ and $\mathbf{q}_F \in \mathbb{R}^{2\times 1}$, respectively.

Problem P1 is difficult to solve due to several obstacles. First, P1 is *nonconvex* due to the nonconvex objective function and constraints C1 and C2. Second, joint optimization of the beamforming vectors $\mathbf{w}_k[n]$ and the UAV's trajectory $\mathbf{q}[n]$ is hindered by their *tight couplings* in (7) and (10), where the steering vector dynamically changes with the UAV's position in a highly nonlinear (nonconvex) manner, cf. (5).

To tackle these challenges, in Sec. IV we reformulate P1 as a multi-stage DP. The reformulated problem can be (asymptotically) optimally solved using the Bellmann optimality equation and SDP, though at the cost of an overwhelming computational complexity. Inspired by the huge success of approximate DP and reinforcement learning, we propose a low-complexity high-quality approximate solution based on the rollout algorithm [13].

IV. Problem Solution

A. Trajectory Design and Rollout Algorithm

In this section, we reformulate P1 as an equivalent N-stage DP problem in a discrete-event dynamic system. We show that P1 can be optimally solved via DP algorithms, by jointly optimizing waypoints $\mathbf{q}[n]$, the beamforming vectors $\mathbf{w}_k[n]$, and the flight durations $\tau[n]$ of problem P1.

1) DP based Reformulation: With a slight abuse of notation, we let n also be the index of stages. The system state $\mathbf{o}_n \triangleq [Q_n, E_n]$ is characterized by the sequence of waypoints $Q_n = \{\mathbf{q}[0], \ldots, \mathbf{q}[n-1]\}$ taken by the UAV till stage n and the UAV's remaining energy E_n at stage n. Meanwhile, the action $\mathbf{z}_n = [\mathbf{q}[n+1], \tau[n+1], \mathbf{w}_k[n+1]]$ consists of the selected next waypoint $\mathbf{q}[n+1] \in \mathcal{V}$, flight duration $\tau[n+1] \in \mathbb{R}^+$ allocated to reach next waypoint, and the beamforming vectors $\mathbf{w}_k[n+1] \in \mathbb{C}^{S\times 1}$ to be used in the next waypoint. The action $\mathbf{z}_n$ is constrained to take values in a given subset $\mathbf{z}_n = \mu_n(\mathbf{o}_n) \in Z(\mathbf{o}_n)$ that depends on the current state $\mathbf{o}_n$. Here, set $Z(\cdot)$ represents the constraints on action $\mathbf{z}_n$ as given in P1. When action $\mathbf{z}_n$ is applied at stage n, the system state evolves to the next state $\mathbf{o}_{n+1}$ according to the system equation $\mathbf{o}_{n+1} = f_n(\mathbf{o}_n, \mathbf{z}_n)$. Meanwhile, a reward $g_n(\mathbf{o}_n, \mathbf{z}_n)$ is generated for the transition from state $\mathbf{o}_n$ to $\mathbf{o}_{n+1}$ under action $\mathbf{z}_n$. For problem P1, we define $g_n(\mathbf{o}_n, \mathbf{z}_n) \triangleq \min_m(\Gamma_m[n+1]-\Gamma_m[n])$, i.e., $g_n(\mathbf{o}_n, \mathbf{z}_n)$ is the increase in the worst-case accumulated sensing energy among all GTs.

For the whole ISAC mission, the actions $\mathbf{z}_n$, $n = 0, \ldots, N-1$, are determined by a policy of the form $\pi = \{\mu_0, \ldots, \mu_{N-1}\}$. Now, given an initial state $\mathbf{o}_0 = [Q_0 = \{\mathbf{q}_I\}, E_0 = E^{\text{tot}}]$ and policy π, the sum of rewards achievable in the ISAC mission is given as $J_\pi(\mathbf{o}_0) = g_N(\mathbf{o}_N) + \sum_{n=0}^{N-1} g_n(\mathbf{o}_n, \mathbf{z}_n)$. Then problem P1 can be equivalently reformulated as seeking the optimal policy π^* that maximizes $J_\pi(\mathbf{o}_0)$ [14], i.e.,

$$\pi^* \in \operatorname*{argmax}_{\pi\in\Pi} J_\pi(\mathbf{o}_0). \tag{14}$$

where set Π includes all policies leading to feasible actions.

2) Bellman Optimality Equation: To solve the multi-stage optimization in (14), let $J_n^*(\mathbf{o}_n)$ be the optimal sum of rewards starting from state $\mathbf{o}_n$ till $\mathbf{o}_N$. $J_n^*(\mathbf{o}_n)$ is also known as the optimal value function of state $\mathbf{o}_n$. Then, the optimal policy can be obtained by solving at each stage $n=0, \ldots, N-1$ the Bellman optimality equation,

$$J_n^*(\mathbf{o}_n) = \max_{\mathbf{z}_n \in Z_n(\mathbf{o}_n)} [g_n(\mathbf{o}_n, \mathbf{z}_n) + J_{n+1}^*(f(\mathbf{o}_n, \mathbf{z}_n))]. \tag{15}$$

In particular, the solution is based on the DP algorithm, which starts from solving the terminal problem $J_N^*(\mathbf{o}_N)$, and then moves one stage backwards to solve $J_N^*(\mathbf{o}_{N-1})$. This process continues until the optimal policy is constructed.

However, executing the DP algorithm (15) is practically infeasible, due to the infinite-space of the continuous action variables. In a simpler case, the actions can be discretized to be specified by a finite set. Yet, the resulting DP algorithm is prohibitively time-consuming except for simple special cases [14]. This motivates us to seek low-complexity approximation methods, such as the rollout algorithm [13], to tackle (15).

Algorithm 1: Joint Trajectory and Beamforming Optimization with OSLR

1: **Input:** $N, \mathbf{r}_d, \tilde{\pi}, E^{\text{tot}}, V_{\max}, \mathbf{q}_I, \mathbf{q}_F, \{\mathbf{u}_k^{\text{c}}, R_{\min,k}\}_{k=1}^{K}, \{\mathbf{u}_m^{\text{s}}\}_{m=1}^{M}$
2: **for** $n = 0 : N-1$ **do** $\triangleright$ For each stage n
3: Find $Z(\mathbf{o}_n)$ based on C1-5; $\triangleright$ Problem P1
4: Calculate $\tilde{J}_{n+1,}(\mathbf{o}_n)$ by simulating the base policy and collecting the rewards;
5: $\mathbf{z}_n^* = \underset{\mathbf{z}_n \in Z(\mathbf{o}_n)}{\operatorname{argmax}} [g_n(\mathbf{o}_n, \mathbf{z}_n) + \tilde{J}_{n+1,}(\mathbf{o}_n)]$;
6: Update system: $\mathbf{o}_{n+1} = f_n(\mathbf{o}_n, \mathbf{z}_n^*)$;
7: **end for**
8: **Output:** $[\mathbf{z}_0, \ldots, \mathbf{z}_{N-1}], J_0^*(\mathbf{o}_0)$.

3) One-Step Lookahead Rollout (OSLR) Algorithm: For the DP algorithm, calculating the optimal value function J_{n+1}^* is computationally intensive. To overcome this difficulty, the OSLR algorithm approximates J_{n+1}^* by the reward function $\tilde{J}_{n+1}$ of a *base policy* $\tilde{\pi} = \{\tilde{\mu}_0, \ldots, \tilde{\mu}_{N-1}\}$. The base policy $\tilde{\pi}$ can be an arbitrary heuristic algorithm for which we can easily calculate its reward function

$$\tilde{J}_{n+1}(\mathbf{o}_{n+1}) \triangleq g_N(\mathbf{o}_N) + \sum\nolimits_{i=n+1}^{N-1} g_i(\mathbf{o}_i, \mathbf{z}_i) \qquad (16)$$

for $\mathbf{z}_i = \tilde{\mu}_i(\mathbf{o}_i), i = n+1, \ldots, N-1$. Using $\tilde{J}_{n+1}$, the action $\tilde{\mathbf{z}}_n$ is selected by applying OSLR as

$$\tilde{\mathbf{z}}_n \in \underset{\mathbf{z}_n \in Z(\mathbf{o}_n)}{\operatorname{argmax}} \left[g_n(\mathbf{o}_n, \mathbf{z}_n) + \tilde{J}_{n+1}(f(\mathbf{o}_n, \tilde{\mu}_n(\mathbf{o}_n))) \right]. \qquad (17)$$

That is, instead of applying the heuristic algorithm itself, the OSLR algorithm optimizes the action by considering both the immediate reward $g_n(\mathbf{o}_n, \mathbf{z}_n)$ in each iteration and the potential long-term reward. Thus, the OSLR algorithm can guarantee a performance improvement over the original base policy with low computational costs [13].

Algorithm 1 summarizes the proposed OSLR algorithm. In each stage, $\tilde{J}_{n+1,}(\mathbf{o}_n)$ is calculated via computationally efficient base policy $\tilde{\pi}$. This requires solving the nonconvex problem P1 for given waypoints as specified by the base policy and system state variable Q_n, cf. line 4. Although the resulting optimization problem is still nonconvex, we will show in Sec. IV-B that it can be efficiently solved to optimality using the SDP relaxation technique. The obtained optimal solutions are further used to calculate the sum of rewards and to select the best action $\mathbf{z}_n$. The state is then updated with the action $\tilde{\mathbf{z}}_n$ that maximizes the accumulated sum reward. States are updated for a total of N stages, as defined in path discretization.

B. Beamforming and Flight Duration Optimization

For Algorithm 1 to work, we still need to optimize the beamforming $\mathbf{w}_k[n]$ and flight duration $\tau[n]$ in the nonconvex problem P1 for given waypoints, i.e. $\mathbf{q}[n]$s, as specified by the base policy and system state variable Q_n. To this end, let $\mathbf{W}_k[n] \triangleq \tau[n]\mathbf{w}_k[n]\mathbf{w}_k^H[n]$, with $\mathbf{W}_k[n] \succeq 0$ and Rank($\mathbf{W}_k[n]$) ≤ 1. Using $\mathbf{W}_k[n]$ as well as defining $\gamma_{\min,k} \triangleq 2^{\frac{R_{\min,k}}{B}} - 1$ and $\mathbf{H}_k[n] \triangleq \mathbf{h}_k[n]\mathbf{h}_k^H[n]$, we can reformulate C1 as

$$\begin{aligned}
\text{C1} \iff & \frac{\tau[n] \cdot |\mathbf{h}_k^H[n]\mathbf{w}_k[n]|^2}{\tau[n] \cdot \left(\sum_{i=1, i\neq k}^{K} |\mathbf{h}_k^H[n]\mathbf{w}_i[n]|^2 + \sigma_k^2\right)} \geq \gamma_{\min,k} \\
\iff & \frac{\text{Tr}(\mathbf{W}_k[n]\mathbf{H}_k[n])}{\sum_{i=1, i\neq k}^{K} \text{Tr}(\mathbf{W}_i[n]\mathbf{H}_k[n]) + \sigma_k^2\tau[n]} \geq \gamma_{\min,k} \\
\iff & \overline{\text{C1}}\text{: } (1 + \gamma_{\min,k}^{-1})\text{Tr}(\mathbf{W}_k[n]\mathbf{H}_k[n]) \\
& - \sum\nolimits_{i=1}^{K} \text{Tr}(\mathbf{W}_i[n]\mathbf{H}_k[n]) \geq \sigma_k^2\tau[n]. \qquad (18)
\end{aligned}$$

Similarly, let $\mathbf{A}_m[n] \triangleq \mathbf{a}_m[n]\mathbf{a}_m^H[n]$. We can rewrite the accumulated sensing energy for GT m as

$$\begin{aligned}
\overline{P^{\text{s}}}_m[n] \cdot \tau[n] &= \sum\nolimits_{k=1}^{K} \text{Tr}(\mathbf{W}_k[n]\mathbf{A}_m[n]), \\
\overline{\Gamma}_m[n] &= \sum\nolimits_{k=1}^{K} \frac{\beta \text{Tr}(\mathbf{W}_k[n]\mathbf{A}_m[n])}{\|\mathbf{d}_m[n]\|^2 + H^2}. \qquad (19)
\end{aligned}$$

Meanwhile, the propulsion energy in (12) is a nonconvex function of speed $V[n]$, which is difficult to optimize. To tackle this challenge, we approximate (12) by a convex function

$$\overline{E}^{\text{f}}(V) \triangleq \left(P_0 \left(1 + \frac{3V^2}{U_{\text{tip}}^2}\right) + \frac{P_i}{1 + \frac{3V}{4V_0}} + \frac{1}{2} d_0 \rho s A V^3 \right) \tau[n], \qquad (20)$$

since $\left(\sqrt{1 + \frac{V^4}{4V_0^4}} - \frac{V^2}{2V_0^2}\right)^{\frac{1}{2}} \approx \left(1 + \frac{3V}{4V_0}\right)^{-1}$ for small $V \geq 0$.

Then, given the waypoints $\mathbf{q}[n]$, P1 can be reduced to and equivalently reformulated as

$$\text{P2}: \max_{\mathbf{W}_k[n], \tau[n]} \ \min_{m} \ \sum\nolimits_{n=1}^{N} \overline{\Gamma}_m[n] \qquad (21)$$

$$\begin{aligned}
s.t.\ \ & \overline{\text{C1}}, \text{C4, C5}, \\
& \overline{\text{C2}}: \sum\nolimits_{n=1}^{N} \left(E^{\text{c}}(\mathbf{w}_k[n]) + \overline{E}^{\text{f}}(V[n])\right) \leq E^{\text{tot}} \\
& \overline{\text{C3}}: \sum\nolimits_{k=1}^{K} \text{Tr}(\mathbf{W}_k[n]) \leq P_{\max}\tau[n], \ \forall n \\
& \text{C6: } \mathbf{W}_k[n] \succeq 0, \forall n, \forall k \qquad (22) \\
& \text{C7: } \text{Rank}(\mathbf{W}_k[n]) \leq 1, \ \forall n, \forall k.
\end{aligned}$$

Now, previously nonconvex constraints in P1, i.e. C1 and C2, are converted to convex constraints $\overline{\text{C1}}$ and $\overline{\text{C2}}$ in Problem P2. However, P2 is still nonconvex due to the rank constraint C7. We solve P2 by employing the SDP relaxation technique, i.e., by dropping constraint C7. The resulting convex optimization problem can be efficiently solved using off-the-shelf solvers such as CVX [15]. Generally, the SDP relaxation approach finds a performance upper bound for the original problem, as the relaxation may enlarge the feasible set. However, for problem P2 at hand, we can show below that the SDP relaxation approach is *tight*, i.e., the relaxed solution always fulfills the rank constraint C7. We have also validated this result offline by simulations.

Lemma 1: Assume that problem P2 admits at least one feasible solution. Then we can always obtain an optimal rank-one solution of $\mathbf{W}_k[n]$ for P2 by solving the relaxed SDP problem.

TABLE I: Parameter Settings for Simulation

Parameter	Notation/Value
Path discretization	$N = 40, \Delta_{\max} = 40$ m
UAV's flight altitude	$H = 100$ m
UAV's max flight speed	$V_{\max} = 20$ m/s
UAV's energy budget	$E^{\text{tot}} = 15000$ J
Number of UAV antennas	$S = 6$
ULA dipole axis vector	$\mathbf{r}_d = [0.7\ 0.7]^T$
Reference channel gain	$\beta = -30$ dB
System bandwidth	$B = 1$ MHz
Noise power	$\sigma_k^2 = -110$ dBm
Maximum transmit power	$P_{\max} = 40$ dBm
Circuitry power consumption	$P_{\text{const}} = 5$ W
Flight power parameters [12]	$A = 0.503$ m^2, $U_{\text{tip}} = 120$m/s, $\rho = 1.225$kg/m, $s = 0.05$m^3, $d_0 = 0.6$, $V_0 = 4.03$, $P_0 = 80$W,$P_i = 88.6$W
Communication requirement	$R_{\min,k} = 2$ Mbits/s

Proof: Please refer to [16, Theorem 2] for a similar proof. ∎

V. Simulation Results

We evaluate the performance the proposed energy-aware joint beamforming and trajectory optimization algorithm for UAV-aided ISAC with dipole antenna array via simulations. We consider a 500m×500m area that is quantized into a uniform grid of 21×21 nodes. In the area, there are $K = 2$ GUs at locations $[126, 101]^T$, $[374, 99]^T$ and $M = 2$ GTs at locations $[126, 376]^T$, $[401, 301]^T$. Given GU/GT locations are identical for all simulations presented. Unless otherwise stated, the simulation parameters are set according to Table I. The UAV's initial and destination locations are set as $\mathbf{q}_I = [50, 50]^T$ and $\mathbf{q}_F = [450, 50]^T$, respectively.

In Algorithm 1, the OSLR requires a *base policy* to approximate the value function (16). For our simulations, we employ a *base policy* for minimizing the distance to the destination location from the UAV position. For performance comparison, we consider two baseline schemes. For baseline scheme 1, we employ the trajectory of the *base policy*, without the improvement brought by the OSLR. For baseline scheme 2, we adopt a UAV equipped with a ULA of isotropic antenna elements, and jointly optimize the ISAC beamforming and trajectory optimization problem with the proposed method.

Figure 3 shows the accumulated sensing energy for the GTs, cf. the objective of P2, versus the communication requirement per GU, $R_{\min}$. We observe that the accumulated sensing energy for all the considered schemes monotonically decreases with increasing $R_{\min}$. This reveals an inherent trade-off between sensing and communication in UAV-aided ISAC. In particular, as $R_{\min}$ increases, the UAV has to fly closer to the GU locations and/or steer the ISAC beam more towards the GUs, to satisfy the more stringent communication constraint, as depicted in Figures 4 and 5. Meanwhile, the proposed scheme significantly outperforms both baseline schemes when $R_{\min} \leq 3$ Mbits/s. For example, compared with the baseline scheme 1 and 2, the proposed scheme increases the accumulated sensing energy by up to a factor of 12 and 2.4, respectively, thanks to the extra DoFs brought by joint trajectory and ISAC beamforming design of directional dipole array.

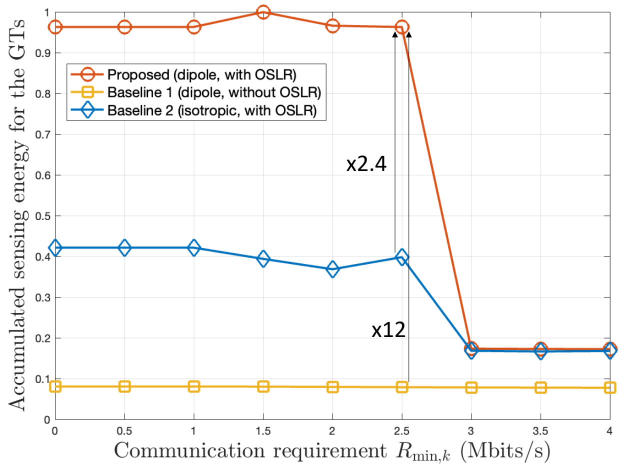

Fig. 3: Accumulated sensing energy for the GTs versus the minimum required data rate per GU.

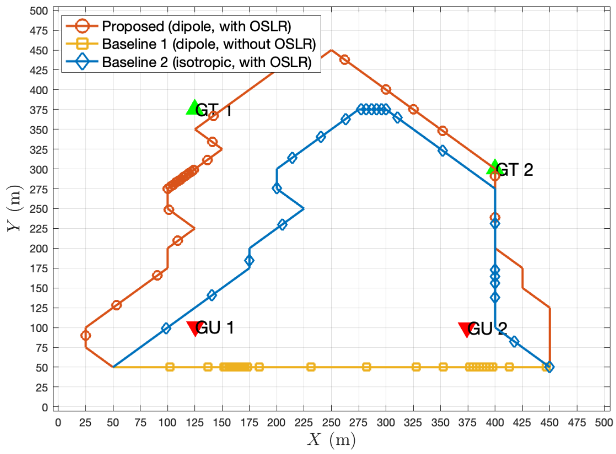

Fig. 4: Optimized flight trajectories and flight durations of the UAV for the considered schemes, $R_{\min,k} = 2$ Mbits/s.

Figures 4 and 5 depict the optimized trajectories and flight durations of the UAV for the proposed and baseline schemes. The two figures correspond to two different communication requirement $R_{\min}$ values. To illustrate flight durations, we mark the trajectories with points evenly spaced in time. That is, each marker point indicates a fixed amount of flight time along the trajectory. We observe that, as expected, the baseline scheme 1 designs the UAV's trajectory using the shortest path between the UAV's initial and destination positions. Unlike the baseline scheme 1, both the proposed scheme and the baseline scheme 2 design the UAV's trajectory to fly toward the GT locations, thanks to the OSLR based Algorithm 1. This can significantly reduce the loss of signal energy for sensing the GTs, while maintaining communication QoS for information transmission to the GUs.

Figure 6 evaluates the accumulated sensing energy for the GTs, versus the number of antenna elements S deployed at

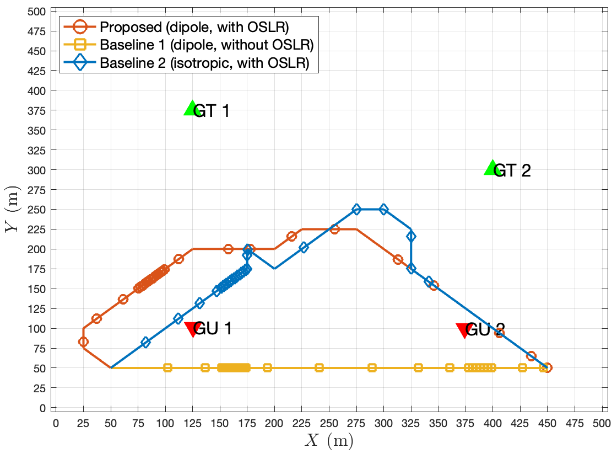

Fig. 5: Optimized flight trajectories and flight durations of the UAV for the considered schemes, $R_{\min,k} = 3$ Mbits/s.

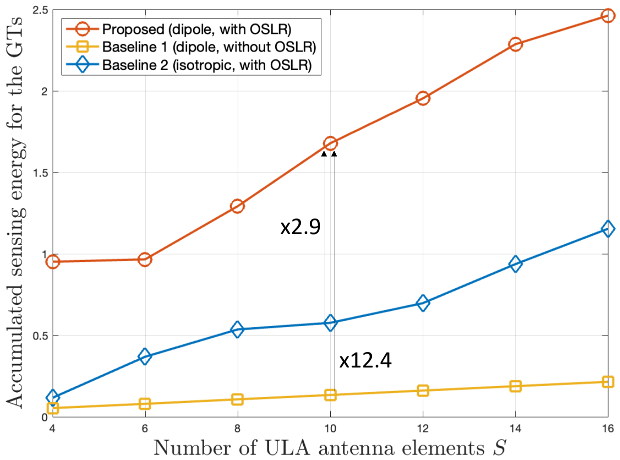

Fig. 6: Accumulated sensing energy for the GTs versus the number of transmit antennas.

the transmit ULA of the UAV. We observe that increasing the number of antennas improves the performances of all schemes, as signal energy can be focused towards the GTs and GUs using narrower beams generated by the antennas. Moreover, our proposed scheme significantly outperforms the baseline schemes 1 and 2, achieving performance gains of up to a factor of 12.4 and 2.9, respectively. This is because the baseline scheme 1 optimizes only the flight times and ISAC beamforming, but limits the flight trajectory to be along the shortest path between the UAV's initial and destination positions. Meanwhile, our proposed scheme jointly optimizes the trajectory design, flight times and ISAC beamforming. Additionally, with the directional antenna array, the proposed scheme can better exploit the UAV trajectory to increase the sensing power toward the GTs while mitigating the inter-user interference than the isotropic antenna array.

VI. Conclusions and Future Work

This paper explored the application of a dipole antenna array for transmitting ISAC signals on an energy-constrained UAV. By considering the 3D radiation pattern of the dipole array, we jointly optimized transmit beamforming using the dipole array and trajectory planning for the UAV during the ISAC mission. Our objective was to maximize the worst-case accumulated sensing energy among all GTs while meeting communication QoS requirements, flight constraints, and energy limitations. To tackle the formulated nonconvex problem, we reformulated it as a multi-stage DP with continuous actions, and subsequently proposed a computationally efficient OSRL algorithm to provide a high-quality suboptimal solution. Simulation results demonstrated that our proposed approach can effectively leverage UAV mobility and spatial DoFs of the transmit dipole array to significantly improve both the sensing and communication performance, achieving up to 12.4 times higher sensing energy compared to two baseline schemes. We recognize that the magnitudes of performance gains can vary with the orientation of the dipole array deployed at the UAV. While we have assumed fixed array orientation in this paper, joint optimization of array orientation, beamforming, and flight trajectory is left as an interesting topic for our future work.

References

[1] F. Liu, Y. Cui *et al.*, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[2] F. Liu, C. Masouros *et al.*, "MU-MIMO communications with MIMO radar: From co-existence to joint transmission," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2755–2770, 2018.

[3] X. Liu, T. Huang *et al.*, "Joint transmit beamforming for multiuser MIMO communications and MIMO radar," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3929–3944, 2020.

[4] M. B. Yilmaz, L. Xiang, and A. Klein, "UAV-Assisted Delay-Sensitive Communications with Uncertain User Locations: A Cost Minimization Approach," in *Proc. IEEE PIMRC*, 2022.

[5] K. Meng, Q. Wu *et al.*, "UAV-Enabled Integrated Sensing and Communication: Opportunities and Challenges," *IEEE Trans. Wireless Commun.*, pp. 1–9, 2023.

[6] K. Zhang and C. Shen, "UAV aided integrated sensing and communications," in *Proc. IEEE VTC*, 2021.

[7] X. Jing, F. Liu, and C. Masouros, "ISAC from the sky: Net-zero energy UAV trajectory design," in *Proc. IEEE ICASSPW*, 2023.

[8] J. Wu, W. Yuan, and L. Hanzo, "When UAVs meet ISAC: Real-time trajectory design for secure communications," *IEEE Trans. Veh. Technol.*, pp. 1–6, 2023.

[9] Z. Lyu, G. Zhu, and J. Xu, "Joint maneuver and beamforming design for UAV-enabled integrated sensing and communication," *IEEE Trans. on Wireless Commun.*, vol. 22, no. 4, pp. 2424–2440, 2023.

[10] I. Orikumhi, J. Bae, and S. Kim, "UAV-assisted integrated sensing and communications for user blockage prediction," in *Proc. IEEE ICTC*, 2022.

[11] A. Khalili, A. Rezaei *et al.*, "Energy-aware resource allocation and trajectory design for UAV-enabled ISAC," *arXiv:2302.10124*, 2023.

[12] Y. Zeng, J. Xu, and R. Zhang, "Energy minimization for wireless communication with rotary-wing UAV," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2329–2345, 2019.

[13] D. P. Bertsekas, *Lessons from AlphaZero for optimal, model predictive, and adaptive control*. Athena Scientific, 2022.

[14] D. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Athena Scientific, 2001, vol. 1 and 2.

[15] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[16] J. Zhang, Y. Zhang *et al.*, "Robust energy-efficient transmission for wireless-powered d2d communication networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7951–7965, 2021.