Mustafa Burak Yilmaz, Lin Xiang, Anja Klein, ”Joint Beamforming and Trajectory Optimization for UAV-Enabled ISAC under a Finite Energy Budget”, in *IEEE International Conference on Communications (ICC)*, Denver, CO, Jun. 2024.

Joint Beamforming and Trajectory Optimization for UAV-Enabled ISAC under a Finite Energy Budget

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 studies joint beamforming and trajectory optimization for integrated sensing and communication (ISAC) enabled by an energy-constrained unmanned aerial vehicle (UAV). The UAV transmits information-bearing signals using an onboard uniform linear array (ULA), for simultaneously serving downlink communication users and sensing multiple targets during its flight. To explore the synergy between UAV and ISAC, we jointly optimize the UAV's flight trajectory, ISAC beamforming, and mission completion time for maximizing the accumulated sensing energy for ground targets under a finite energy budget for the UAV, while guaranteeing quality-of-service for communication users. The formulated problem is highly non-convex, which is generally intractable. Motivated by the success of approximate dynamic programming (DP) methods, we propose a novel low-complexity high-quality solution by combining the one-step lookahead rollout algorithm in approximate DP and the semidefinite programming technique in convex optimization. Our simulation results show that, compared to two baseline schemes, the proposed scheme can significantly enlarge the achievable sensing and communication performance region for ISAC.**

I. Introduction

Integrated sensing and communication (ISAC) is an advanced physical layer technology designed to enable simultaneous sensing and communication in the sixth-generation (6G) wireless networks, utilizing common transmit signals over shared spectrum and transmitter hardware [1]. Recently, enabling ISAC via unmanned aerial vehicles (UAVs) has attracted a growing research interest. The flexible mobility of UAVs can significantly enhance the performance of ISAC and extend the range of ISAC services [2], [3], particularly in areas of disrupted or limited network infrastructures, in both normal and emergency situations [4]. So far, UAV-enabled ISAC has been widely considered for simultaneous collection and transmission of sensed data [5], joint sensing/localization and communication [3], [6], as well as sensing-/tracking-assisted (secure) communication [7].

However, UAVs are usually constrained by size, weight, and power (SWAP). To facilitate efficient UAV-aided ISAC in 6G, several recent works have explored joint optimization of UAV movements and beamforming design for ISAC signal transmission in UAV-enabled ISAC systems [8]–[10]. In [8], the authors optimized UAV's hovering position/flight trajectory and transmit beamforming for maximizing the communication data rate while ensuring sufficient beam pattern gain on sensing targets. The authors of [9] exploited sensed information regarding the propagation environment to optimize the selection of beamforming vectors from a given codebook, aiming to maximize the communication data rate in a downlink multiuser scenario. However, UAV's flight energy was ignored in both [8] and [9]. In [10], the authors optimized transmit beamforming, UAV trajectory, and sensing schedule to minimize the UAV's total energy consumption over a given mission completion time, with sensing performed at fixed locations. However, predetermining a "suitable" mission completion time in [10] is a nontrivial task, which should be sufficient for the mission while avoiding energy waste, for energy-efficient operations. Moreover, the previous works [8]–[10] underscore difficulties in solving the joint beamforming and trajectory optimization problem for UAV-enabled ISAC, a formidable endeavor even when the UAV's flight energy is ignored. Existing optimization methods and exhaustive search approaches employed in tackling the problem in [8]–[10] can be either inefficient or computationally intensive.

To bridge these research gaps, this paper introduces a novel energy-aware joint beamforming and trajectory design for UAV-enabled ISAC. The UAV, equipped with a uniform linear array (ULA), is tasked with transmitting information-bearing signals to simultaneously serve downlink communication users and sense multiple targets during its flight. In order to best exploit the UAV with limited energy, we jointly optimize the beamforming, trajectory, and the mission time to maximize the sensing performance while satisfying the quality-of-service (QoS) requirements of the communication users and respecting a total energy budget for the UAV. Compared with fixing the mission time a priori as in [10], optimizing it under the energy budget avoids energy waste and can significantly improve energy utilization. Our contributions are:

- We investigate joint trajectory design, transmit beamforming, and mission completion time optimization for a UAV-enabled ISAC system with finite onboard energy. The aim is to maximize the accumulated sensing performance over the whole mission while satisfying communication, flight and energy requirements.
- The formulated problem is highly nonconvex. By reformulating it as a multi-stage dynamic program (DP), we propose a novel efficient solution via combining the one-step lookahead rollout (OSLR) technique in approximate DP and the semidefinite programming (SDP) in convex optimization.
- Simulation results show that the proposed OSLR algorithm can jointly exploit the UAV mobility and the multi-antenna beamforming to significantly outperform two baseline schemes in enlarging the achievable sensing and communication performance region for ISAC.

Throughout this paper, matrices and vectors are denoted by boldface capital and lower-case letters, respectively. $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathrm{Rank}(\mathbf{A})$, and $\mathrm{Tr}(\mathbf{A})$ denote the transpose, Hermitian conjugate transpose, rank, and trace of matrix $\mathbf{A}$, respectively. $|\cdot|$ and $\|\cdot\|$ denote the absolute value of a complex scalar and the Euclidean norm of a complex vector, respectively. Finally, $\frac{d}{dx}$ denotes the first-order derivative with respect to x.

II. System Model

We consider a UAV-enabled ISAC system as illustrated in Fig. 1. A rotary-wing UAV equipped with a ULA of $S \geq 1$

This work has been funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center and has been supported by the BMBF project Open6GHub under grant 16KISKO14 and by DAAD with funds from the German Federal Ministry of Education and Research (BMBF).

979-8-3503-0405-3/24/$31.00 ©2024 IEEE

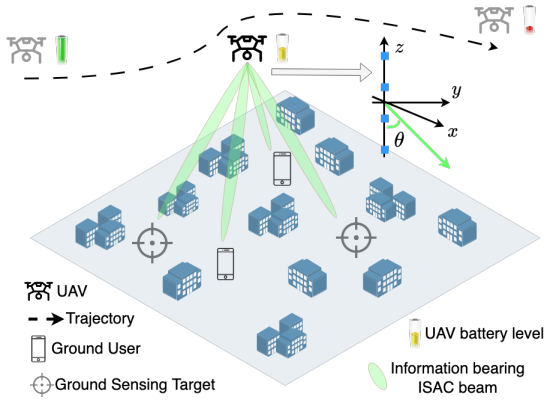

Fig. 1. System model of UAV-enabled ISAC with limited onboard battery.

antennas is deployed as a dual-functional access point to simultaneously communicate with K single-antenna ground users (GUs) and sense M ground targets (GTs). Following [8], [10], we assume that the ULA is vertically placed, cf. Fig. 1. As such, the GUs and GTs located at different distances from the UAV observe distinct angles of departure (AoDs) at the ULA. The GU k, $k = 1, ..., K$, and GT m, $m = 1, ..., M$, are assumed to be at fixed locations given by $\mathbf{u}_k^{\mathrm{c}} \in \mathbb{R}^{2\times 1}$ and $\mathbf{u}_m^{\mathrm{s}} \in \mathbb{R}^{2\times 1}$, respectively. For convenience of presentation, we consider a continuous-time system model with time index t throughout this section and formulate a continuous-time optimization problem in Sec. III. We then seek an approximate solution of the formulated problem in Sec. IV, by discretizing the UAV's flight path. We assume that the UAV flies at a fixed altitude H during its mission, whereas its horizontal position $\mathbf{q}(t) \in \mathbb{R}^{2\times 1}$ can vary flexibly over time. To avoid depletion of the battery, the UAV needs to complete the mission within a finite energy budget E^{tot}.

A. Channel and Signal Models of UAV-Enabled ISAC

We assume that the UAV's flying altitude H is sufficiently large such that a strong LoS link exists between the UAV and each GU [2]. Let $\mathbf{h}_k(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})$ be the channel vector from the UAV to GU k at time t. We model $\mathbf{h}_k(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})$ as

$$\mathbf{h}_k(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}}) = \frac{\sqrt{\beta}}{d(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})} \cdot \mathbf{a}(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}}), \tag{1}$$

where β denotes the channel power gain at a unit distance and $d(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}}) \triangleq \sqrt{\|\mathbf{q}(t) - \mathbf{u}_k^{\mathrm{c}}\|^2 + H^2}$ is the Euclidean distance between the UAV and GU k at time t. Moreover, $\mathbf{a}(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})$ is the steering vector of the ULA toward GU k at time t,

$$\mathbf{a}(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}}) \triangleq [1, e^{jz(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})}, \ldots, e^{jz(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})(S-1)}]^T, \tag{2}$$

where we define $z(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}}) \triangleq 2\pi \frac{d^{\mathrm{ULA}}}{\lambda} \cos(\theta(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}}))$ with spacing d^{ULA} between adjacent ULA elements and carrier wavelength λ. Here, due to the vertical array placement, the steering vector depends only on the AoD $\theta(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})$ of GU k observed from the UAV, which is given as

$$\theta(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}}) \triangleq \arccos\left(\frac{H}{\sqrt{\|\mathbf{q}(t) - \mathbf{u}_k^{\mathrm{c}}\|^2 + H^2}}\right). \tag{3}$$

Let $c_k \in \mathbb{C}$ and $\mathbf{w}_k \in \mathbb{C}^{S\times 1}$ be the information-bearing symbol intended for GU k and its associated beamforming vector, respectively. We assume that c_k is a circularly symmetric complex Gaussian (CSCG) random variable with zero mean and unit variance. The UAV's transmit signal is given by $\mathbf{x}(t) = \sum_{i=1}^{K} \mathbf{w}_i(t)\, c_i(t)$. Meanwhile, the received signal $\mathbf{y}_k(t)$ at GU k at time t is given by

$$\mathbf{y}_k(t) = \mathbf{h}_k^H(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})\, \mathbf{x}(t) + n_k(t), \tag{4}$$

where $n_k(t)$ is the additive white Gaussian noise at GU k and is modeled as a zero-mean Gaussian random variable with variance σ_k^2. Assume a system bandwidth of B Hz. Based on (4), the instantaneous data rate of GU k at time t is given as

$$R_k(t) = B \log_2\left(1 + \frac{|\mathbf{h}_k^H(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})\mathbf{w}_k(t)|^2}{\sum_{i=1, i\neq k}^{K} |\mathbf{h}_k^H(\mathbf{q}(t), \mathbf{u}_k^{\mathrm{c}})\mathbf{w}_i(t)|^2 + \sigma_k^2}\right). \tag{5}$$

We consider radar sensing towards the GTs, by reusing the communication signal $\mathbf{x}(t)$. To this end, the UAV collects echos of the communication signal reflected/scattered from the GTs, which are further used to e.g. detect the presence of active users, track changes in the target area [8], or separate the GTs by e.g. their angles of arrival and round-trip times at the radar receiver [1]. For these applications, the objective is to improve the *accumulated sensing energy* for the GTs, namely the total sensing energy radiated to the GTs during the whole mission. This metric can also be extended to evaluate the energy of reflected/scattered echo signals when received at the UAV with matched filters, by including additional power losses from the GTs to the UAV. However, our problem formulation and solution remain applicable for both metrics. Assume that, similar to communications, the sensing is performed over a LoS-dominant link with path loss model as in (1). The accumulated sensing energy for GT m is given as

$$\Gamma_m(T) = \int_{t=0}^{T} \frac{\beta P^{\mathrm{s}}(\mathbf{q}(t), \mathbf{u}_m^{\mathrm{s}})}{d^2(\mathbf{q}(t), \mathbf{u}_m^{\mathrm{s}})}\, dt, \tag{6}$$

where T is the overall mission time and its value depends on the energy utilization within the budget during ISAC and flight. Moreover, $P^{\mathrm{s}}(\mathbf{q}(t), \mathbf{u}_m^{\mathrm{s}})$ is the power of transmitted signal in the direction of $\mathbf{u}_m^{\mathrm{s}}$ at time t, given as

$$P^{\mathrm{s}}(\mathbf{q}(t), \mathbf{u}_m^{\mathrm{s}}) = \sum_{k=1}^{K} |\mathbf{w}_k^H(t)\mathbf{a}(\mathbf{q}(t), \mathbf{u}_m^{\mathrm{s}})|^2. \tag{7}$$

B. Energy Consumption of the UAV

The UAV consumes energy during both ISAC transmission and flight. The instantaneous power consumption for signal transmission at time t, denoted by $P^{\mathrm{c}}(\mathbf{w}_k(t))$, is given as

$$P^{\mathrm{c}}(\mathbf{w}_k(t)) = \sum_{k=1}^{K} \mathbf{w}_k^H(t)\mathbf{w}_k(t) + P_{\mathrm{const}}, \tag{8}$$

which includes the transmit power in the first term and a constant power, P_{const}, consumed in the circuitry and signal processing etc. Let $V(t) = \|d\,\mathbf{q}(t)/dt\|$ be the instantaneous speed of the UAV. The instantaneous propulsion power consumption of the rotary-wing UAV is modeled as [11]

$$P^{\mathrm{f}}(V(t)) = P_0\left(1 + \frac{3V(t)^2}{U_{\mathrm{tip}}^2}\right) + P_i\left(\sqrt{1 + \frac{V(t)^4}{4V_0^4}} - \frac{V(t)^2}{2V_0^2}\right)^{1/2} + \frac{1}{2} d_0 \rho s A V(t)^3. \tag{9}$$

TABLE I
PARAMETERS OF FLIGHT POWER CONSUMPTION MODEL (9).

Notation/Value	Definition
$\Omega = 300$ rad/s	Blade angular velocity
$A = 0.503$ m^2	Rotor disc area
$U_{\text{tip}} = 120$ m/s	Tip speed of the rotor blade
$s = 0.05$ m^3	Rotor solidity
$\rho = 1.225$ kg/m^3	Air density
$d_0 = 0.6$	Fuselage drag ratio
$V_0 = 4.03$ m/s	Mean rotor velocity induced in forward flight
$P_0 = 80$ W	Blade profile power during hovering
$P_i = 88.6$ W	Induced power during hovering

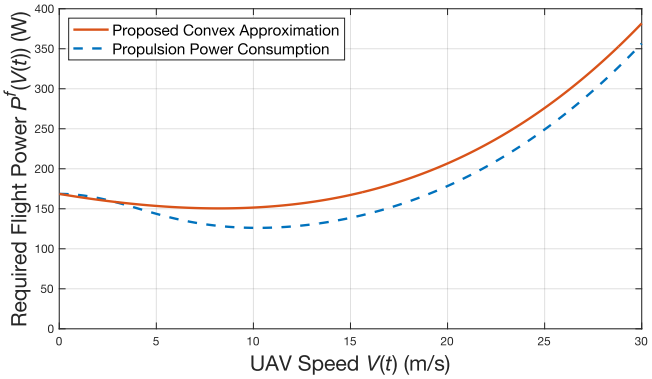

Fig. 2. Comparison of propulsion power function (9) and its approximation (10), using typical parameter settings as shown in Table I.

Table I summarizes the definition and typical values of the parameters involved in (9) [11].

However, the second term in (9), for modeling the power consumption during hovering, is a nonconvex function of speed $V(t)$. So is the propulsion power $P^{\text{f}}(V(t))$, which is difficult to optimize. To tackle this challenge, we approximate the nonconvex term in (9) by a convex function

$$\overline{P}^{\text{f}}(V) \triangleq P_0\left(1+\frac{3V^2}{U_{\text{tip}}^2}\right)+P_i\left(\frac{1}{1+3\frac{V}{4V_0}}\right)+\frac{1}{2}d_0\rho sAV^3, \quad (10)$$

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 $V \geq 0$. A comparison in Fig. 2 shows that (9) and (10) closely match with each other for a wide range of flight speeds. Note that (9) can also be (successively) approximated by other convex functions for improved accuracy, where our proposed optimization solution still applies, though at increased computational costs.

III. Problem Formulation

This paper aims to jointly exploit the spatial degrees of freedom (DoFs) of the transmit ULA and the mobility of the UAV to maximize the performance gains of UAV-enabled ISAC within the given energy budget. In order to properly illuminate the GTs, we use the accumulated sensing energy defined in (6) as the sensing performance metric, similar to [10]. Moreover, assuming that GU and GT locations are known, we formulate the joint trajectory, beamforming, and mission time optimization problem for UAV-enabled ISAC under finite energy budget as

$$\text{P1}: \max_{\mathbf{q}(t),\mathbf{w}_k(t),T} \quad \sum\nolimits_{m=1}^{M} \Gamma_m \quad (11)$$
$$\text{s.t. C1a: } R_k(t) \geq R_{\min,k} \ \forall k, \forall t \in [0,T]$$
$$\text{C1b: } |\Gamma_m - \Gamma_o| \leq E_{\text{diff}}, \forall m \neq o$$
$$\text{C1c: } \int_{t=0}^{T} \left(P^{\text{c}}(\mathbf{w}_k(t)) + \overline{P}^{\text{f}}(V(t))\right) dt \leq E^{\text{tot}}$$
$$\text{C1d: } V(t) \leq V_{\max}, \forall t \in [0,T]$$
$$\text{C1e: } \sum\nolimits_{k=1}^{K} \mathbf{w}_k^H(t)\mathbf{w}_k(t) \leq P_{\max}, \forall t \in [0,T]$$
$$\text{C1f: } \mathbf{q}(0) = \mathbf{q}_I, \quad \mathbf{q}(T) = \mathbf{q}_F.$$

In problem P1, constraint C1a ensures that the UAV maintains communication with GU k during the ISAC mission with an instantaneous rate exceeding $R_{\min,k}$ to guarantee communication QoS. C1b requires similar levels of accumulated sensing energy for different GTs to enhance fairness while sensing the GTs [12]. C1c defines a finite total energy budget for the UAV mission. C1d and C1e limit the maximum flight speed and the maximum transmit power of the UAV by $V_{\max}$ and $P_{\max}$, respectively. Finally, C1f specifies the UAV's starting position and destination as $\mathbf{q}_I \in \mathbb{R}^{2\times 1}$ and $\mathbf{q}_F \in \mathbb{R}^{2\times 1}$, respectively.

Problem P1 poses several challenges for solution. Firstly, P1 is *nonconvex* due to its nonconvex objective function and constraints C1a, C1b, and C1c. Secondly, the objective function and all constraints in P1 depend on the *free* variable T, which is not given a priori, but must be determined via optimization. Thirdly, the beamforming vectors $\mathbf{w}_k(t)$ and the UAV's trajectory $\mathbf{q}(t)$ are *tightly coupled* in (4) and (7), where the steering vector dynamically changes with the UAV's position in a highly nonlinear and nonconvex manner, cf. (2). Lastly, both $\mathbf{q}(t)$ and $\mathbf{w}_k(t)$ are continuous functions over time t, involving an *infinite* number of variables, and P1 contains continuous-time integrals, which are inconvenient to handle.

To tackle these challenges, in Sec. IV we discretize the UAV's flight path and reformulate P1 as a multi-stage DP. The reformulated problem can be (asymptotically) optimally solved using the Bellman equation and SDP, with the structure of the optimal solution also revealed, though at an overwhelming computational complexity [13]. Inspired by the huge success of approximate DP and reinforcement learning, we further propose a low-complexity high-quality approximate solution based on the rollout algorithm [14].

IV. Problem Solution

A. Path Discretization

We approximate the continuous-time optimization problem P1 by discretizing the UAV's flight path as in [11]. Compared with the well-known time discretization method [8], [10], the considered path discretization seems more suitable for ISAC application, as it can drastically reduce the number of optimization variables when UAV hovers or flies at low speeds along the trajectory. Here, the UAV's flight path is discretized into N line segments defined by $N+1$ waypoints $\{\mathbf{q}_n\}_{n=0}^{N}$, where $\mathbf{q}_0 = \mathbf{q}_I$, $\mathbf{q}_N = \mathbf{q}_F$, and

$$||\mathbf{q}_{n+1} - \mathbf{q}_n|| \leq \Delta_{\max}, \forall n. \quad (12)$$

We require $\Delta_{\max} \ll H$ such that the distances $d(\mathbf{q}(t), \mathbf{u}_k^{\text{c}})$ and $d(\mathbf{q}(t), \mathbf{u}_m^{\text{s}})$ between the UAV and each GU/GT rarely change when the UAV flies along each line segment. Note that the total length of the line segments, $N\Delta_{\max}$ gives a lower bound on the total flight distance. This bound tightens as N increases, but extremely large N should be avoided for computational savings. Moreover, assume that the UAV flies at a constant speed on each line segment, with the flight time over the nth line segment denoted by τ_n. Once the values of τ_ns

are determined, we can calculate the UAV's flying speed along line segment n as $V_n = ||\mathbf{q}_{n+1}-\mathbf{q}_n||/\tau_n$ and the mission time as $T = \sum_{n=1}^{N} \tau_n$. Hence, problem P1 can be approximated as

$$\begin{aligned}
\text{P2}: \max_{\mathbf{q}_n, \mathbf{w}_{n,k}, \tau_n} \quad & \sum\nolimits_{m=1}^{M} \Gamma_{N,m} \\
\text{s.t.} \quad & \text{C2a: } R_{n,k} \geq R_{\min,k} \ \forall k, n \\
& \text{C2b: } |\Gamma_{N,m} - \Gamma_{N,o}| \leq E_{\text{diff}}, \forall m \neq o \\
& \text{C2c: } \sum\nolimits_{n=1}^{N} \left(P_n^{\text{c}} + \overline{P}^{\text{f}}(V_n) \right) \tau_n \leq E^{\text{tot}} \\
& \text{C2d: } \sum\nolimits_{k=1}^{K} \mathbf{w}_{n,k}^H \mathbf{w}_{n,k} \leq P_{\max}, \forall n \\
& \text{C2e: } ||\mathbf{q}_{n+1} - \mathbf{q}_n|| \leq \min(\Delta_{\max}, V_{\max}\tau_n), \forall n \\
& \text{C2f: } \mathbf{q}_0 = \mathbf{q}_I, \quad \mathbf{q}_N = \mathbf{q}_F,
\end{aligned}$$

where $\Gamma_{N,m}$ is the accumulated sensing energy for GT m, cf. (6), over the N line segments and is given as

$$\Gamma_{N,m} = \sum\nolimits_{n=1}^{N} \frac{\beta P_{n,m}^{\text{s}}}{d^2(\mathbf{q}_n, \mathbf{u}_m^{\text{s}})} \tau_n. \tag{13}$$

$R_{n,k}$ is the instantaneous rate of GU k while the UAV flies over line segment n and is defined as

$$R_{n,k} = B \log_2 \left(1 + \frac{|\mathbf{h}_k^H(\mathbf{q}_n, \mathbf{u}_k^{\text{c}}) \mathbf{w}_{n,k}|^2}{\sum_{i \neq k} |\mathbf{h}_k^H(\mathbf{q}_n, \mathbf{u}_k^{\text{c}}) \mathbf{w}_{n,i}|^2 + \sigma_k^2} \right). \tag{14}$$

Other notations in P2 can be derived from their continuous-time counterparts in the same manner.

B. Beamforming and Flight Duration Optimization

Assume that the waypoints $\mathbf{q}_n$s are given. We show below that $\mathbf{w}_{n,k}$ and τ_n in problem P2 can be jointly optimized using the SDP approach. Although SDP has been used for optimal beamforming design in [8]–[10], we consider here a different problem setting. In Sec. IV-C, we further extend the SDP approach to tackle the joint optimization problem P2.

Let $\mathbf{W}_{n,k} \triangleq \tau_n \mathbf{w}_{n,k} \mathbf{w}_{n,k}^H$, with $\mathbf{W}_{n,k} \succeq 0$ and $\text{Rank}(\mathbf{W}_{n,k}) \leq 1$. Using $\mathbf{W}_{n,k}$ as well as defining $\gamma_{\min,k} \triangleq 2^{\frac{R_{\min,k}}{B}} - 1$ and $\mathbf{H}_{n,k} \triangleq \mathbf{h}_k(\mathbf{q}_n, \mathbf{u}_k^{\text{c}}) \mathbf{h}_k^H(\mathbf{q}_n, \mathbf{u}_k^{\text{c}})$, we can reformulate C2a as

$$\begin{aligned}
\text{C2a} \iff & \frac{\tau_n \cdot |\mathbf{h}_k^H(\mathbf{q}_n, \mathbf{u}_k^{\text{c}}) \mathbf{w}_{n,k}|^2}{\tau_n \cdot \left(\sum_{i=1, i \neq k}^{K} |\mathbf{h}_{n,k}^H \mathbf{w}_{n,i}|^2 + \sigma_k^2 \right)} \geq \gamma_{\min,k} \\
\iff & \frac{\text{Tr}(\mathbf{W}_{n,k} \mathbf{H}_{n,k})}{\sum_{i=1, i \neq k}^{K} \text{Tr}(\mathbf{W}_{n,i} \mathbf{H}_{n,k}) + \sigma_k^2 \tau_n} \geq \gamma_{\min,k} \\
\iff & \text{C3a: } (1 + \gamma_{\min,k}^{-1}) \text{Tr}(\mathbf{W}_{n,k} \mathbf{H}_{n,k}) \\
& \qquad - \sum\nolimits_{i=1}^{K} \text{Tr}(\mathbf{W}_{n,i} \mathbf{H}_{n,k}) \geq \sigma_k^2 \tau_n.
\end{aligned} \tag{15}$$

Then, P2 can be reformulated as

$$\begin{aligned}
\text{P3}: \max_{\mathbf{W}_{n,k}, \tau_n} \quad & \sum\nolimits_{m=1}^{M} \overline{\Gamma}_{N,m} \\
\text{s.t.} \quad & \text{C3a, C2c, C2e, C2f,} \\
& \text{C3b: } \left| \overline{\Gamma}_{N,m} - \overline{\Gamma}_{N,o} \right| \leq E_{\text{diff}}, \forall m \neq o \\
& \text{C3d: } \sum\nolimits_{k=1}^{K} \text{Tr}(\mathbf{W}_{n,k}) \leq P_{\max} \tau_n, \ \forall n \\
& \text{C3g: } \mathbf{W}_{n,k} \succeq 0, \forall n, \forall k
\end{aligned} \tag{16}$$

$$\text{C3h: } \text{Rank}(\mathbf{W}_{n,k}) \leq 1, \ \forall n, \forall k.$$

Where, we similarly let $\mathbf{A}_{n,m} \triangleq \mathbf{a}_m(\mathbf{q}_n, \mathbf{u}_m^{\text{s}}) \mathbf{a}_m^H(\mathbf{q}_n, \mathbf{u}_m^{\text{s}})$ and rewrite the accumulated sensing energy for GT m as

$$\begin{aligned}
\overline{P^{\text{s}}}_{n,m} \cdot \tau_n &= \sum\nolimits_{k=1}^{K} \text{Tr}(\mathbf{W}_{n,k} \mathbf{A}_{n,m}), \\
\overline{\Gamma}_{N,m} &= \sum\nolimits_{n=1}^{N} \sum\nolimits_{k=1}^{K} \frac{\beta \text{Tr}(\mathbf{W}_{n,k} \mathbf{A}_{n,m})}{d^2(\mathbf{q}_n, \mathbf{u}_m^{\text{s}})}.
\end{aligned} \tag{17}$$

Problems P3 and P2 are *equivalent* in the sense that they achieve exactly the same optimal objective value and their optimal solutions bear a one-to-one correspondence.

However, P3 is still nonconvex due to the nonconvex rank constraint C3h. We solve P3 by employing the SDP relaxation technique, i.e., by dropping C3h. 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 relaxed solution of $\mathbf{W}_{n,k}$ may violate C3h. However, for problem P3 at hand, we can show below (as well as offline in simulations) that the SDP relaxation approach is *tight*, i.e., the relaxed solution always fulfills C3h.

Lemma 1: Assume that the relaxed SDP problem P3 admits at least one feasible solution. Then an optimal rank-one solution of $\mathbf{W}_{n,k}$ can always be obtained by solving P3.

Proof: Our proof follows a similar approach as [16, Theorem 2], but extends it to solve the joint beamforming and flight duration optimization problem P2 for UAV-aided ISAC. Due to limited page space, the details are ignored. ■

C. Trajectory Design and Rollout Algorithm

1) DP based Reformulation: Based on the results in Sec. IV-B, we now proceed to jointly optimize the waypoints $\mathbf{q}_n$, the beamforming vectors $\mathbf{w}_{n,k}$, and the flight durations τ_n, to solve problem P2. To this end, we first reformulate P2 as an N-stage DP problem in a discrete-event dynamic system and then apply the DP algorithm for solution. In particular, let n be the index of stages. The system is represented by state $\mathbf{o}_n \triangleq [Q_n, E_n]$, which includes the sequence of waypoints $Q_n = \{\mathbf{q}_0, \ldots, \mathbf{q}_n\}$ known till stage n, and the remaining energy E_n. Due to the system dynamics, the state evolves according to the equation $\mathbf{o}_{n+1} = f_n(\mathbf{o}_n, \mathbf{z}_n)$, when action $\mathbf{z}_n$ is applied at stage n, and f_n is a function of $(\mathbf{o}_n, \mathbf{z}_n)$ that describes the mechanism by which the state is updated from stage n to $n+1$. The action $\mathbf{z}_n = [\mathbf{q}_{n+1}, \tau_{n+1}, \mathbf{w}_{n+1,k}]$ consists of the selected next waypoint $\mathbf{q}_{n+1} \in \mathbb{R}^{2\times 1}$, flight duration $\tau_{n+1} \in \mathbb{R}^+$ allocated to reach next waypoint, and the beamforming vectors $\mathbf{w}_{n+1,k} \in \mathbb{C}^{S\times 1}$ to be used in the next waypoint. The action $\mathbf{z}_n$ is determined by a policy of the form $\pi = \{\mu_0, \ldots, \mu_{N-1}\}$, and 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 $Z(\cdot)$ captures all constraints on action $\mathbf{z}_n$ given in P2. We denote by Π the set of all policies constructed using feasible actions.

The transition from state $\mathbf{o}_n$ to $\mathbf{o}_{n+1}$ under the action $\mathbf{z}_n$ generates a reward $g_n(\mathbf{o}_n, \mathbf{z}_n)$. We define $g_n(\mathbf{o}_n, \mathbf{z}_n)$ as the incremental sensing energy for the GTs at stage n, i.e., $g_n(\mathbf{o}_n, \mathbf{z}_n) \triangleq \sum_{m=1}^{M} \frac{\beta P_{n+1,m}^{\text{s}}}{d^2(\mathbf{q}_{n+1}, \mathbf{u}_m^{\text{s}})} \tau_{n+1}$. Now, given an initial state $\mathbf{o}_0 = [Q_0 = \{\mathbf{q}_I\}, E_0 = E^{\text{tot}}]$ and policy π, the sum of the rewards, denoted as $J_\pi(\mathbf{o}_0) = g_N(\mathbf{o}_N) + \sum_{n=0}^{N-1} g_n(\mathbf{o}_n, \mathbf{z}_n)$, recaps the accumulated sensing energy. Hence, problem P2 is

equivalent to seeking from the set Π the optimal policy π^* that maximizes $J_\pi(\mathbf{o}_0)$ [13], i.e.,

$$\pi^* \in \operatorname{argmax}_{\pi\in\Pi} J_\pi(\mathbf{o}_0). \tag{18}$$

The resulting optimal sum of rewards starting from state $\mathbf{o}_n$ till $\mathbf{o}_N$, denoted by $J_n^*(\mathbf{o}_n)$, is known as the optimal value function of state $\mathbf{o}_n$.

Note that (18) involves a multi-stage optimization. To obtain the optimal policy, we need to solve at each stage $n = 0,\ldots,N-1$ the Bellman optimality equation,

$$J_n^*(\mathbf{o}_n) = \max_{\mathbf{z}_n\in Z_n(\mathbf{o}_n)} [g_n(\mathbf{o}_n,\mathbf{z}_n) + J_{n+1}^*(f_n(\mathbf{o}_n,\mathbf{z}_n))], \tag{19}$$

using e.g. the DP algorithm. We start with solving the terminal problem $J_N^*(\mathbf{o}_N)$, and then move one stage backwards to solve $J_{N-1}^*(\mathbf{o}_{N-1})$. This process continues until $J_0^*(\mathbf{o}_0)$ is obtained. Based on $J_n^*(\mathbf{o}_n)$, $n = 0,\ldots,N-1$, the optimal policy is further constructed via forward optimization steps, see [13] for details. However, executing the DP algorithm (19) is prohibitively time-consuming, due to the infinite-space of the continuous action variables (even if they are discretized into a finite set of numbers) [13]. This motivates us to seek practical low-complexity approximation methods, such as the rollout algorithm [14], to tackle (19).

2) One-Step Lookahead Rollout (OSLR) Algorithm: For the DP algorithm, the bottleneck for computation lies in calculating the optimal value function $J_{n+1}^*(\cdot)$. To overcome this difficulty, the OSLR algorithm approximates $J_{n+1}^*(\cdot)$ by the reward function $\tilde{J}_{n+1}(\cdot)$ of a *base policy* $\tilde{\pi}=\{\tilde{\mu}_0,\ldots,\tilde{\mu}_{N-1}\}$. The base policy $\tilde{\pi}$ is a heuristic algorithm whose reward $\tilde{J}_{n+1}(\mathbf{o}_{n+1}) \triangleq g_N(\mathbf{o}_N) + \sum_{i=n+1}^{N-1} g_i(\mathbf{o}_i,\mathbf{z}_i)$ can be conveniently calculated for $\mathbf{z}_i=\tilde{\mu}_i(\mathbf{o}_i), i = n+1,\ldots,N-1$. Using $\tilde{J}_{n+1}(\cdot)$, the action $\tilde{\mathbf{z}}_n$ for stage n is selected by OSLR as

$$\tilde{\mathbf{z}}_n \in \operatorname*{argmax}_{\mathbf{z}_n\in Z(\mathbf{o}_n)} [g_n(\mathbf{o}_n,\mathbf{z}_n) + \tilde{J}_{n+1}(f(\mathbf{o}_n,\tilde{\mu}_n(\mathbf{o}_n))]. \tag{20}$$

Algorithm 1 summarizes the proposed OSLR algorithm. In each stage n, $\tilde{J}_{n+1}(\mathbf{o}_n)$ is calculated via the computationally efficient base policy $\tilde{\pi}$. Based on the sequence of selected actions (waypoints) $\tilde{\mathbf{z}}_0,\ldots,\tilde{\mathbf{z}}_{n-1}$ and actions $\mathbf{z}_{n+1},\ldots,\mathbf{z}_{N-1}$ deduced by base policy $\tilde{\pi}$, we utilize the formulated SDP problem P3 to calculate the sum rewards of multiple actions to select the best action $\tilde{\mathbf{z}}_n$ (20). The state is then updated using the action $\tilde{\mathbf{z}}_n$ that maximizes the accumulated sum reward. Procedure followed for all N stages, to find all waypoints. Let $\mathcal{O}(\Omega)$ be the complexity of utilized heuristic algorithm in base policy $\tilde{\pi}$, and $|Z|$ be the maximum number of actions at a stage. Then the overall complexity of Algorithm 1 is bounded by $\mathcal{O}(N|Z|\Omega)$, much lower than that of the DP algorithm. Moreover, instead of applying the heuristic algorithm itself, the OSLR algorithm optimizes the action by considering the potential long-term reward (20). As such, the OSLR algorithm guarantees a performance improvement over the original base policy with low computational costs [14].

Algorithm 1 Joint Trajectory and Beamforming Optimization with OSLR

Input: $N, \tilde{\pi}, E^{\mathrm{tot}}, V_{\max}, \mathbf{q}_I, \mathbf{q}_F, \{\mathbf{u}_k^{\mathrm{c}}, R_{\min,k}\}_{k=1}^K, \{\mathbf{u}_m^{\mathrm{s}}\}_{m=1}^M$;
for $n = 0 : N-1$ **do** ▷ For each stage n
  Find $Z(\mathbf{o}_n)$ based on C2a-f; ▷ P2
  Calculate $\tilde{J}_{n+1,}(\mathbf{o}_n)$ by simulating the base policy and collecting the rewards;
  $\mathbf{z}_n^* = \operatorname*{argmax}_{\mathbf{z}_n\in Z(\mathbf{o}_n)} [g_n(\mathbf{o}_n,\mathbf{z}_n) + \tilde{J}_{n+1,}(\mathbf{o}_n)]$;
  Update system: $\mathbf{o}_{n+1} = f_n(\mathbf{o}_n,\mathbf{z}_n^*)$;
end for
Output: $[\mathbf{z}_0,\ldots,\mathbf{z}_{N-1}], J_0^*(\mathbf{o}_0)$.

V. Simulation results

We evaluate the performance the proposed joint beamforming and trajectory optimization algorithm for UAV-enabled ISAC with finite energy budget via simulations. We consider a $500\text{m}\times500\text{m}$ area that is divided into 400 smaller squares with size $25\text{m}\times25\text{m}$. This results in 441 grid nodes to be considered as UAV waypoints. In the area there are $K = 2$ GUs and $M = 2$ GTs, with locations shown in Fig. 3. Unless otherwise stated, the simulation parameters are set according to Table I and II. The UAV's initial and destination locations are set as $\mathbf{q}_I = [50, 50]^T$ and $\mathbf{q}_F = [50, 450]^T$, respectively.

TABLE II
Parameter Settings for Simulation

Parameter	Notation/Value
Path discretization length	$\Delta_{\max} = 35$ m
Number of UAV antennas	$S = 6$
UAV's max flight speed	$V_{\max} = 20$ m/s
UAV's flight altitude	$H = 100$ m
UAV's energy budget	$E^{\mathrm{tot}} = 15000$ J
Fairness constraint, cf. C3b	$E_{\mathrm{diff}} = 0.1$
Channel gain at reference distance	$\beta = -30$ dB
Noise power	$\sigma_k^2 = -110$ dBm
Maximum transmit power	$P_{\max} = 40$ dBm
Communication requirement per GU	$R_{\min,k} = 2$ Mbps
System bandwidth	$B = 1$ MHz
Circuitry power consumption	$P_{\mathrm{const}} = 5$ W

Recall that OSLR employs a *base policy* to approximate the value function in Algorithm 1. In this paper, 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 OSLR. For baseline scheme 2, we adopt zero-forcing (ZF) beamforming for mitigating interference between the GUs, and jointly optimize flight durations and trajectory similar to problem P2.

Figure 3 depicts the optimized trajectories and flight durations of the UAV for the proposed and baseline schemes. 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. As GT 1 is located further away from the designed trajectory than GT 2, the UAV flies near GT 1 for a long time, indicated by dense time marks on the trajectory. This ensure fairness for sensing, cf. C2b in problem P2. Unlike the baseline scheme 1, both the proposed scheme and the baseline scheme 2 design the UAV's trajectories 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. Recall that the base policy of the OSLR algorithm is the same as the baseline scheme 1. This result implies that the OSLR algorithm can significantly outperform the base policy, by

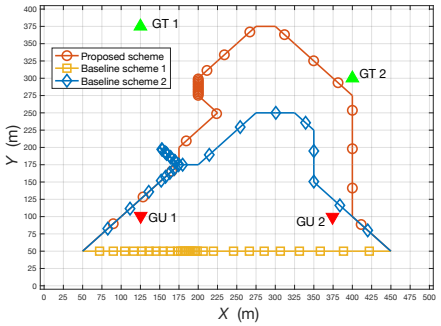

Fig. 3. Optimized flight trajectories and flight durations of the UAV for the considered schemes.

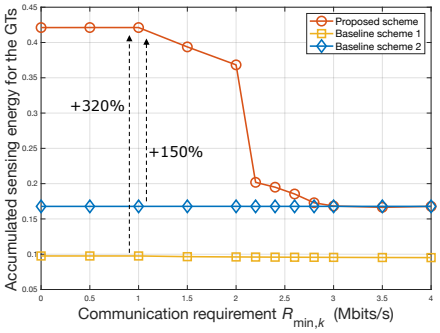

Fig. 4. Accumulated sensing energy for the GTs versus the minimum required data rate per GU.

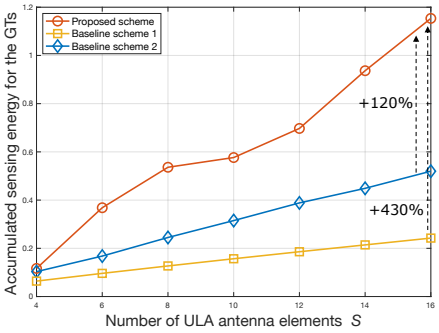

Fig. 5. Accumulated sensing energy for the GTs versus the number of transmit antennas.

using the latter to simulate and approximate future long-term rewards in (20). However, constrained by ZF beamforming, the baseline scheme 2 is bounded to remain close to the GUs at later parts of the trajectory. On the other hand, the proposed scheme flexibly adapts the ISAC to continue meeting the communication requirements while flying closer to the GTs.

Figure 4 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 decrease with $R_{\min}$. This reveals an inherent trade-off between sensing and communication in UAV-enabled 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, cf. C1a. Meanwhile, the proposed scheme significantly outperforms both baseline schemes when $R_{\min} \leq 2$ Mbits/s. For example, compared with the baseline scheme 1 and 2, the proposed scheme increases the accumulated sensing energy by up to 320% and 150%, respectively, thanks to the extra DoFs brought by joint trajectory and ISAC beamforming design. Interestingly, the proposed scheme achieves the same performance as the baseline scheme 2 when $R_{\min} \geq 3$ Mbits/s. This result suggests that for UAV-enabled ISAC, ZF beamforming may be optimal when $R_{\min}$ is large and communication requirements dominate the beamforming and trajectory design.

Finally, Figure 5 evaluates the accumulated sensing energy for the GTs versus the number of antennas S deployed deployed at 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, claiming gains of up to 430% and 120%, 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, the baseline scheme 2 optimizes only flight times and trajectory by assuming ZF beamforming. In contrast, our proposed scheme jointly optimizes the trajectory design, flight times and ISAC beamforming. Hence, with the proposed scheme, the UAV can better exploit the antenna array to increase the sensing power toward the GTs while mitigating the inter-user interference.

VI. Conclusions and Future Work

In this paper we studied joint transmit beamforming optimization and trajectory design for UAV-enabled ISAC with limited onboard energy. We formulated a highly nonconvex optimization problem to maximize sensing performance while satisfying QoS requirements of communication, as well as flight and energy constraints. By discretizing the flight path and reformulating the problem as a multi-stage DP with continuous states and actions, we proposed a computationally efficient OSRL algorithm to obtain a high-quality suboptimal solution. Simulation results revealed that the proposed scheme can jointly exploit the UAV mobility and the spatial DoFs of transmit ULA to significantly enlarge the achievable sensing and communication performance region. The performance gains are up to 430% compared to two baseline schemes. Motivated by its high performance and low complexity, extending the proposed OSLR algorithm to other base policies and/or multi-UAV enabled ISAC are interesting topics left for our future work.

References

[1] F. Liu *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] K. Meng *et al.*, "UAV-Enabled Integrated Sensing and Communication: Opportunities and Challenges," *IEEE Trans. Wireless Commun.*, pp. 1–9, 2023.
[3] F. Pei *et al.*, "Joint optimization of beamforming and 3D array-steering for UAV-aided ISAC," in *Proc. IEEE ICC*, 2024(Accepted).
[4] M. B. Yilmaz *et al.*, "UAV-Assisted Delay-Sensitive Communications with Uncertain User Locations: A Cost Minimization Approach," in *Proc. IEEE PIMRC*, 2022.
[5] K. Zhang and C. Shen, "UAV aided integrated sensing and communications," in *Proc. IEEE VTC*, 2021.
[6] X. Jing *et al.*, "ISAC from the sky: Net-zero energy UAV trajectory design," in *Proc. IEEE ICASSPW*, 2023.
[7] J. Wu *et al.*, "When UAVs meet ISAC: Real-time trajectory design for secure communications," *IEEE Trans. Veh. Technol.*, pp. 1–6, 2023.
[8] Z. Lyu *et al.*, "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.
[9] I. Orikumhi *et al.*, "UAV-assisted integrated sensing and communications for user blockage prediction," in *Proc. IEEE ICTC*, 2022.
[10] A. Khalili *et al.*, "Energy-aware resource allocation and trajectory design for UAV-enabled ISAC," *Proc. IEEE Globecom*, 2023.
[11] Y. Zeng *et al.*, "Energy minimization for wireless communication with rotary-wing UAV," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2329–2345, 2019.
[12] Z. Wang *et al.*, "NOMA empowered integrated sensing and communication," *IEEE Commun.Lett.*, vol. 26, no. 3, pp. 677–681, 2022.
[13] D. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Athena Scientific, 2001, vol. 1 and 2.
[14] D. P. Bertsekas, *Lessons from AlphaZero for optimal, model predictive, and adaptive control*. Athena Scientific, 2022.
[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 *et al.*, "Robust energy-efficient transmission for wireless-powered d2d communication networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7951–7965, 2021.