Yi Wang, Lin Xiang, and Anja Klein, “Energy-Constrained UAV Trajectory Design for Sensing Information Maximization” accepted for presentation in *IEEE Global Communications Conference (Globecom)*, Kuala Lumpur, Malaysia, December 2023.

Energy-Constrained UAV Trajectory Design for Sensing Information Maximization

Yi Wang, Lin Xiang, and Anja Klein
Communications Engineering Lab, Technische Universität Darmstadt, Germany
Emails: {y.wang, l.xiang, a.klein}@nt.tu-darmstadt.de

***Abstract*—This paper explores the use of an unmanned aerial vehicle (UAV) equipped with sensing devices to collect data, such as temperature and wind speed, in a target area. We distinguish from prior studies by considering *correlation* in the data sensed at different positions. The correlation is captured by a spatial autoregressive (SAR) model with distance-dependent covariance matrix. Assuming that the covariance matrix is given, we optimize the UAV's trajectory under energy constraints to maximize the information, namely the joint (differential) entropy, contained in the sensed data. The formulated problem is a mixed-integer nonconvex problem, which is generally intractable. To tackle this challenge, we reformulate the problem to an equivalent dynamic programming (DP) problem and further solve it by a low-complexity trajectory planning scheme based on the One-Step Lookahead Rollout (OSLR) algorithm. Our simulation results under diverse covariance matrices, generated synthetically and by using real-world data, show that the proposed trajectory planning scheme can effectively exploit the UAV's mobility and the spatial correlation of the sensed data. Furthermore, it significantly improves the performance by up to 38% compared to a conventional greedy search-based trajectory design.**

I. Introduction

Unmanned aerial vehicles (UAVs) have been used to collect data from ground sensors or to perform aerial onboard sensing in versatile applications, including crop sensing in precision agriculture, terrain surveying, and meteorological monitoring [1], [2]. UAV-aided sensing has the advantages of achieving high accuracy, large coverage, high adaptability to the environment, and fast task completion. Compared with sensing from satellites and ground sensor networks, UAV-aided sensing has significantly lower deployment costs, requiring neither installing numerous ground sensors nor launching expensive satellites [1]. Moreover, UAVs carrying different types of sensors can be dispatched to the target area and perform multiple tasks simultaneously. Therefore, UAV-aided sensing plays a vital role for enabling integrated space-air-and-ground sensing in the sixth-generation (6G) wireless networks.

However, due to the continuity and similarity of the surrounding environments, the data from different locations are often *correlated*, leading to redundancy and providing no additional information. Moreover, as UAVs usually have a severely limited energy supply, sensing redundant data leads to unnecessary flight and waste of flight time/energy, which unfavorably degrades the performance of UAV-aided sensing. To facilitate efficient UAV-aided sensing, the UAV's trajectory has to be planned carefully to maximize the information in the sensed data subject to the energy supply constraint.

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.

The problem induced by correlated sensed data has also been encountered in terrestrial wireless sensor networks (WSNs), where sensor placement optimization has proven to be an effective technique for increasing the information in the sensed data [3], [4]. Therein, inspired by information theory, entropy-based metrics have been used to measure the information in sensed data and further for optimizing the sensor placement. In [3], the authors employed mutual information to maximize the amount information about the unexplored area one can acquire through the sensed data. However, the metric mutual information considered in [3] constitutes only a part of the total information gathered by the placed sensors. By focusing solely on mutual information for sensor placement, the self-information of the data gathered by the sensors is ignored. On the other hand, unlike mutual information, joint entropy covers the total information of the data gathered from the sensors. The authors of [4] investigated the joint entropy maximization problem for the sensor placement and showed the importance of the sensor placement in enhancing the information in the sensed data. Analogous to sensor placement, strategic trajectory planning is crucial in UAV-aided sensing.

So far, UAVs have traditionally served as data collection platforms for terrestrial WSNs [5]–[7]. In [5], [6], the UAV's trajectory was optimized to minimize the energy consumption of the terrestrial WSNs. In [7], the association between multiple UAVs and sensor nodes (SNs) was optimized by maximizing the entropy to facilitate efficient data collection from terrestrial SNs. However, when UAVs are only utilized as data collectors for terrestrial SNs, the inherent flexibility and ease of deployment for advanced sensing are underutilized. To leverage these advantages, recent studies have also explored the use of UAVs as aerial SNs that actively sense and collect data [8], [9], and investigated the related trajectory planning techniques under energy constraints. In [8], multi-UAV trajectory planning was studied to ensure obstacle-aware ground coverage with minimized flight energy. In [9], the trajectory of an energy-constrained UAV is optimized to maximize the quality of its captured images. The correlation among the captured images is taken into account by considering the diminishing returns of more images sensed for one target. Note that the sensed data's correlation and information were neglected in [7], [8]. The diminishing returns of more images for one target in [9] under-represent the correlation and the information in the sensed data.

Unlike previous studies [5]–[9], this paper models the spatially correlated data in a slowly time-varying environment using a spatial autoregressive (SAR) model. The SAR is widely used to model ecological, economic and demographic

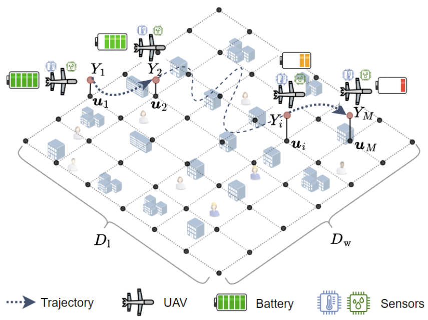


Fig. 1. Sensing with an energy-constrained fixed-wing UAV, which flies over a path before depleting the energy allocated for the sensing task.

data with spatial autocorrelations [10]. We then employ joint (differential) entropy from information theory to quantify the information within the correlated sensed data. Moreover, we consider a trajectory planning scheme for an energy-constrained UAV to maximize the information contained in the sensed data. To our knowledge, no existing literature has yet reported on entropy-based trajectory optimization for UAV-aided sensing that considers the correlation between the data. Our contributions can be summarized as follows:

- We deploy an energy-constrained UAV equipped with onboard sensors to sense correlated data from a target area. The trajectory of the UAV is optimized to maximize the information in the correlated sensed data, using the joint entropy as a metric to quantify the information.
- The formulated optimization problem is a mixed-integer nonconvex problem, which is generally intractable. To this end, we reformulate the problem as a dynamic programming (DP) problem. Subsequently, we propose a novel trajectory planning scheme based on the One-Step Lookahead Rollout (OSLR) algorithm, which offers a computationally efficient suboptimal solution to the DP problem.
- The simulation results show that the proposed OSLR-based solution outperforms the baseline greedy search scheme in terms of achieving a higher joint entropy with limited flight energy or time. Moreover, trajectories obtained from the proposed scheme prioritize the exploration of the area with more information and can also ensure more thorough area coverage.

In the remainder of this paper, Section II presents the system model of UAV-aided sensing. The formulation and the solution of the joint entropy maximization problem are provided in Sections III and IV, respectively. Section V evaluates the performance of the proposed trajectory planning scheme. Finally, Section VI concludes the paper.

II. System Model

In this section, we present the system model for UAV-aided sensing and the SAR model of correlated sensed data. We also introduce the joint entropy to measure the information within the correlated sensed data.

A. UAV-aided Sensing

As depicted in Figure 1, a fixed-wing UAV equipped with a sensor is considered for performing aerial sensing tasks, such as temperature and wind monitoring. Thereby, the UAV is deployed to gather sensed data from an area of width D_w and length D_l. We assume that the UAV flies at a constant altitude during data collection. Moreover, the UAV is allocated a maximum amount of energy, $E_{\max}$, for data collection. Here we assume that the UAV has sufficient amount of energy left for flying back after completing the data collection. Moreover, for miniature sensors, the power consumption is significantly smaller compared with the flight power consumption. Therefore, we neglect the power consumption for sensing.

To facilitate trajectory planning, we consider a discrete-time system, where each time slot has duration $\Delta \geq 0$. Let i be the index of time slots. Moreover, the target area is equally divided into small rectangles formed by N grid nodes located in set $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_N\}$, cf. Figure 1. Let $\boldsymbol{K}_i = [\boldsymbol{u}_0, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_i]$ be the vector of the locations of the grid nodes selected to sense data for the UAV until time slot i, which represents the trajectory of the UAV, with $\boldsymbol{u}_0$ serving as the starting point of the UAV. In each time slot, the UAV can sense data at any grid node under the speed constraints between the minimum speed $V_{\min}$ and the maximum speed and $V_{\max}$.

B. Modeling of Correlated Sensed Data

In this paper, we model the correlated sensed data using the SAR. To this end, let $\boldsymbol{Z} = [Z_1, ..., Z_N]^T \in \mathbb{R}^{N\times 1}$ be a random vector representing the sensed data at the grid nodes within $\mathcal{V}$, where Z_n is the random variable representing the sensed data at grid node n located at $\mathbf{v}_n$. Analogous to the autoregressive models in time series [10], the SAR model of $\boldsymbol{Z}$ is given as

$$\boldsymbol{Z} = \rho \boldsymbol{W}\boldsymbol{Z} + \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{1}$$

where $\rho \in [0, 1]$ is the spatial lag parameter and reflects the overall spatial autocorrelation of the data in the whole area. Moreover, the weight matrix $\boldsymbol{W} \in \mathbb{R}^{N\times N}$ represents the location-specific spatial relation of the sensed data. The diagonal elements are set to zero $w_{n,n} = 0, n = 1, ..., N$ to exclude the self-neighbor relation, and the off-diagonal elements are given by $w_{n,n'} = \exp(-\theta \cdot d_{n,n'})$ for $n \neq n'$, $n' = 1, ..., N$, where $\theta > 0$ denotes the rate of decreasing in spatial autocorrelation with the distance $d_{n,n'}$ between two sensing locations. The weight matrix $\boldsymbol{W}$ is symmetric, i.e., $\boldsymbol{W} = \boldsymbol{W}^T$, since $d_{n,n'} = d_{n',n}$. Moreover, $\boldsymbol{A} \in \mathbb{R}^{N\times L}$ contains independent input variables that affect the output vector $\boldsymbol{Z}$. For example, $\boldsymbol{Z}$ may denote the temperature in a certain area, and $A_{n,l}$ can then be the latitude or local population density of this area. $\boldsymbol{\beta} \in \mathbb{R}^{L\times 1}$ is a weight vector for $\boldsymbol{A}$. Finally, $\boldsymbol{\epsilon} = [\epsilon_1, ..., \epsilon_N]^T \in \mathbb{R}^{N\times 1}$ represents the noise vector and is modeled as a zero-mean Gaussian random vector with covariance matrix $\sigma^2 \boldsymbol{I}_N$.

Based on (1), the data value at a specific location is influenced by the nearby data, with the extent of this influence determined by both ρ and $\boldsymbol{W}$. If $(\boldsymbol{I} - \rho\boldsymbol{W})$ is invertible, $\boldsymbol{Z}$ in (1) can be rewritten as

$$\boldsymbol{Z} = (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}(\boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{\epsilon}). \tag{2}$$

In our paper, we assume that ρ, $\boldsymbol{W}$, $\boldsymbol{A}$ and $\boldsymbol{\beta}$ are known a priori. Therefore, $\boldsymbol{Z}$ follows a multivariate normal distribution, $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where the mean $\boldsymbol{\mu}$ is given by $E[\boldsymbol{Z}] = (\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \boldsymbol{A}\boldsymbol{\beta}$ and covariance matrix is given by

$$\begin{aligned}\boldsymbol{\Sigma} = E[\boldsymbol{Z}\boldsymbol{Z}^T] =& \sigma^2 (\boldsymbol{I} - \rho \boldsymbol{W})^{-1} (\boldsymbol{I} - \rho \boldsymbol{W}^T)^{-1} \\ &+ (\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \boldsymbol{A}\boldsymbol{\beta}\boldsymbol{\beta}^T \boldsymbol{A}^T (\boldsymbol{I} - \rho \boldsymbol{W}^T)^{-1}.\end{aligned} \tag{3}$$

C. Joint Entropy based Measure of Sensed Information

We use joint differential entropy, or simply joint entropy, to quantify the information in the data that the UAV collects. Let $\boldsymbol{Y}_i = [Y_0, Y_1, \cdots, Y_i]$ be the vector of random variables corresponding to data samples collected at locations $\boldsymbol{K}_i = [\boldsymbol{u}_0, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_i]$. Note that the random variables in $\boldsymbol{Y}_i$ are selected (possibly with repetition) from $\boldsymbol{Z}$. Joint (differential) entropy of multiple random variables $\boldsymbol{Y}_i$ is defined as

$$H(\boldsymbol{Y}_i) = -E\left[\log f(y_0, \cdots, y_i)\right], \tag{4}$$

where $f(y_0, \cdots, y_i)$ is the joint probability density function.

The joint entropy $H(\boldsymbol{Y}_i)$ of the random variables $\boldsymbol{Y}_i$ can also be derived using the chain rule

$$H(\boldsymbol{Y}_i) = H(Y_0) + H(Y_1|\boldsymbol{Y}_0) + \cdots + H(Y_i|\boldsymbol{Y}_{i-1}). \tag{5}$$

The conditional entropy of $Y_j, j \leq i$, given $\boldsymbol{Y}_{j-1}$ is given using its conditional variance $\sigma^2_{Y_j|\boldsymbol{Y}_{j-1}}$ [3]

$$H(Y_j|\boldsymbol{Y}_{j-1}) = \frac{1}{2}\log(2\pi e \sigma^2_{Y_j|\boldsymbol{Y}_{j-1}}), \tag{6}$$

where the conditional variance is given by [3]

$$\sigma^2_{Y_j|\boldsymbol{Y}_{j-1}} = \sigma^2_{Y_j} - \boldsymbol{c}_{Y_j\boldsymbol{Y}_{j-1}} \boldsymbol{C}^{-1}_{\boldsymbol{Y}_{j-1}\boldsymbol{Y}_{j-1}} \boldsymbol{c}^T_{Y_j\boldsymbol{Y}_{j-1}}, \tag{7}$$

where $\sigma^2_{Y_j}$ denotes the variance of Y_j, $\boldsymbol{c}_{Y_j\boldsymbol{Y}_{j-1}} = [c_{Y_jY_1}, c_{Y_jY_2}, \cdots, c_{Y_jY_{j-1}}]$ is the vector that contains the values of covariance between Y_j and $\boldsymbol{Y}_{j-1}$. $\boldsymbol{C}^{-1}_{\boldsymbol{Y}_{j-1}\boldsymbol{Y}_{j-1}}$ is the inverse matrix of the covariance matrix of the random variables $\boldsymbol{Y}_{j-1}$. For the special case where $\boldsymbol{Y}_{j-1}$ contains Y_j, $H(Y_j|\boldsymbol{Y}_{j-1}) = 0$.

III. Problem Formulation

Based on Section II, the trajectory of the UAV, along which the UAV collects the sensed data, determines the amount of information contained in the sensed data. In this section, we formulate a trajectory optimization problem to maximize the information contained in the sensed data collected by the UAV with limited energy $E_{\max}$. To this end, we maximize the joint entropy $H(\boldsymbol{Y}_M)$ of the random variables $\boldsymbol{Y}_M$, where M represents the total flight time until the UAV's energy is depleted. Let $\boldsymbol{K}_M = [\boldsymbol{u}_0, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_M]$ denote the trajectory of the UAV. Given the UAV's starting point $\boldsymbol{u}_0$, minimum and maximum velocity $V_{\min}$ and $V_{\max}$, and the covariance matrix $\boldsymbol{\Sigma}$, the resulting problem is formulated as

$$\mathcal{P}_1: \max_{\boldsymbol{u}_i \in \mathcal{V}, M \in \mathbb{Z}_+} \quad H(\boldsymbol{Y}_M) \tag{8}$$

$$\text{subject to: } \mathrm{C}_1: \sum_{i=1}^{M} P_i(v_i)\cdot\Delta \leq E_{\max}, \tag{8a}$$

$$\mathrm{C}_2: V_{\min} \leq v_i \leq V_{\max}, \quad 1 \leq i \leq M. \tag{8b}$$

Constraint C_1 limits the maximum flight energy to $E_{\max}$. The power consumption is given by $P_i(v_i) = \alpha_1 v_i^3 + \frac{\alpha_2}{v_i}$, where $v_i = \|\boldsymbol{u}_i - \boldsymbol{u}_{i-1}\|/\Delta$ is the speed of the UAV in time slot i, and $\alpha_1, \alpha_2 > 0$ are the parameters of the flight energy consumption model of the UAV. The flight speed v_i of the UAV in time slot i is restricted by the minimum and maximum flight speeds $V_{\min}$ and $V_{\max}$ in constraint C_2.

$\mathcal{P}_1$ is a mixed-integer nonconvex problem since it involves the integer variable M and the objective function is nonconvex w.r.t. the selection of the M grid nodes for sensing data. Such type of problem is generally difficult to be optimally solved within a polynomial computation time.

IV. Problem Solution

In this section, we first show that $\mathcal{P}_1$ can be reformulated into an equivalent DP problem and optimally solved using DP algorithms. However, the resulting computational complexity is overwhelming. Inspired by recent success in approximate DP and reinforcement learning, we propose a low-complexity suboptimal solution based on the OSLR algorithm, which tackles the problem by approximating the optimal reward function in the Bellman optimality equation.

A. Dynamic Programming based Reformulation

Let $\boldsymbol{x}_i = [\boldsymbol{u}_i, \boldsymbol{K}_{i-1}, E_i]$ be the system state in time slot i, including the UAV's current location $\boldsymbol{u}_i$, previously visited locations $\boldsymbol{K}_{i-1}$, and the UAV's remaining flight energy E_i in time slot i. Accordingly, we model the system state updating equation by

$$\boldsymbol{x}_{i+1} = f(\boldsymbol{x}_i, \boldsymbol{s}_i), \quad i = 0, \cdots, M, \tag{9}$$

where $\boldsymbol{s}_i$ is the action, i.e., the UAV's movement at time i. $\boldsymbol{s}_i \in S_i(\boldsymbol{x}_i) = \{\boldsymbol{s}_i \in \mathbb{R}^2 \mid v_{i,\min} \leq \frac{\|\boldsymbol{s}_i\|}{\Delta} \leq v_{i,\max}, \boldsymbol{u}_i + \boldsymbol{s}_i \in \mathcal{V}\}$. $v_{i,\min}$ and $v_{i,\max}$ denote the lower and upper bounds on the UAV's speed in time slot i, respectively. $v_{i,\min} \in \{\mathbb{R} \mid v_{i,\min} \geq V_{\min}, P_i(v_{i,\min})\Delta \leq E_i\}$, and $v_{i,\max} \in \{\mathbb{R} \mid v_{i,\max} \leq V_{\max}, P_i(v_{i,\max})\Delta \leq E_i\}$, which guarantees that $v_{i,\min}$ and $v_{i,\max}$ does not surpass the UAV's minimum and maximum flight speeds and the resulting flight energy consumption does not exceed the remaining flight energy E_i in time slot i. The action $\boldsymbol{s}_i$ is determined by the policy $\boldsymbol{s}_i = \pi_i(\boldsymbol{x}_i)$. The policy space Π_i comprises all functions $\pi_i(\cdot): \boldsymbol{x}_i \rightarrow \boldsymbol{s}_i$ that map a state $\boldsymbol{x}_i$ to an action $\boldsymbol{s}_i \in S(\boldsymbol{x}_i)$. Moreover, $\boldsymbol{x}_M$ is the terminal state when $S_M(\boldsymbol{x}_M) = \emptyset$.

The system state updating equation $f(\cdot)$ is defined as

$$\begin{cases}\boldsymbol{u}_{i+1} = \boldsymbol{u}_i + \boldsymbol{s}_i, \\ \boldsymbol{K}_i = [\boldsymbol{K}_{i-1}, \boldsymbol{u}_i], \\ E_{i+1} = E_i - P_{i+1}(v_{i+1})\Delta. \end{cases} \tag{10}$$

Let $g(\boldsymbol{x}_i, \boldsymbol{s}_i)$ represent the reward resulting from the transition from state $\boldsymbol{x}_i$ to $\boldsymbol{x}_{i+1}$ when action $\boldsymbol{s}_i$ is taken. We define $g(\boldsymbol{x}_i, \boldsymbol{s}_i) = H(Y_{i+1}|\boldsymbol{Y}_i)$, which represents the increase in the joint entropy by collecting the sensed data at location $\boldsymbol{u}_{i+1}$. Based on (5), we have $H(\boldsymbol{Y}_M) = \sum_{i=0}^{M} H(Y_i|\boldsymbol{Y}_{i-1}) = H(Y_0) + \sum_{i=0}^{M-1} g(\boldsymbol{x}_i, \boldsymbol{s}_i)$, i.e., the sum of the reward gives the information in the sensed data collected at the grid nodes $\{\boldsymbol{u}_0, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_M\}$ by the UAV.

Now let $J_{\boldsymbol{s}}(\boldsymbol{x}_0) = H(Y_0) + \sum_{i=0}^{M-1} g(\boldsymbol{x}_i, \boldsymbol{s}_i)$ denote the sum of the reward after taking the sequence of actions $\boldsymbol{s} = \{\boldsymbol{s}_0, \cdots, \boldsymbol{s}_{M-1}\}$ starting from the initial state $\boldsymbol{x}_0$. Then problem $\mathcal{P}_1$ can be solved via finding the optimal actions $\boldsymbol{s}^*$ for the following optimization problem [12]

$$\boldsymbol{s}^* = \underset{\boldsymbol{s}}{\operatorname{argmax}}\, J_{\boldsymbol{s}}(\boldsymbol{x}_0). \tag{11}$$

This requires solving the following Bellman optimality equation at each time slot $i = 0, 1, \ldots, M-1$,

$$J_i^*(\boldsymbol{x}_i) = \max_{\boldsymbol{s}_i \in S_i(\boldsymbol{x}_i)} [g(\boldsymbol{x}_i, \boldsymbol{s}_i) + J_{i+1}^*(f(\boldsymbol{x}_i, \boldsymbol{s}_i))], \tag{12}$$

where $J_i^*(\boldsymbol{x}_i)$ denotes the optimal sum reward starting from state $\boldsymbol{x}_i$ to $\boldsymbol{x}_M$, and the action $\boldsymbol{s}_i$ is selected in time slot i to maximize the sum reward $J_i^*(\boldsymbol{x}_i)$. Note that $J_M^*(\boldsymbol{x}_M) = 0$, since the action space $S_M(\boldsymbol{x}_M) = \emptyset$.

To solve the Bellman equation (12), the DP algorithm calculates the optimal sum reward $J_{M-1}^*(\boldsymbol{x}_{M-1}), \cdots, J_0^*(\boldsymbol{x}_0)$ by starting from the terminal state $\boldsymbol{x}_M$ and going backwards in a recursive manner, until the optimal sum rewards starting at all states, i.e., $J_{M-1}^*(\boldsymbol{x}_{M-1}), \cdots, J_0^*(\boldsymbol{x}_0)$, are obtained. Then, a forward algorithm is further used to attain the optimal actions $\{\boldsymbol{s}_0^*, \boldsymbol{s}_1^*, \cdots, \boldsymbol{s}_{M-1}^*\}$, and states $\{\boldsymbol{x}_0, \boldsymbol{x}_1^*, \cdots, \boldsymbol{x}_{M-1}^*\}$ with the given initial state $\boldsymbol{x}_0$. Please refer to [13, Sec. 6.4] for details of the DP solution for deterministic finite horizon problems.

However, the DP algorithm attains the optimal solution for problem $\mathcal{P}_1$ at the cost of an overwhelming computational complexity. Particularly, when the number of states M and/or the number of actions become large, DP suffers from the curse of dimensionality. Moreover, such a complex solution is not applicable for the UAV-aided sensing since the UAV only has limited energy and computing resources.

B. Proposed Suboptimal Solution based on OSLR Algorithm

Here we propose a low-complexity suboptimal solution based on the OSLR algorithm to overcome the shortcomings of DP. The rollout algorithm has been successfully employed in AlphaGo [14] and our problem of trajectory planing for the UAV also resembles to some extent placing the chess pieces on the board. The difference lies in that here we apply rollout to solve an optimization, rather than learning, problem.

To motivate the rollout algorithm, let us start with a heuristic policy $\pi_{\mathrm{h}}(\cdot)$ which chooses the action $\boldsymbol{s}_i$ that maximizes the reward $g(\boldsymbol{x}_i, \boldsymbol{s}_i)$ in time slot i, so called greedy search. The problem is given as follows

$$\mathcal{P}_3 : \max_{\boldsymbol{s}_i \in S_i(\boldsymbol{x}_i)} \quad g(\boldsymbol{x}_i, \boldsymbol{s}_i). \tag{13}$$

Note that $\mathcal{P}_3$ only considers the instantaneous reward of taking action $\boldsymbol{s}_i$ at state $\boldsymbol{x}_i$, but completely ignores $J_{i+1}^*(f(\boldsymbol{x}_i, \boldsymbol{s}_i))$ in subproblem (12). As such, applying the heuristic policy in each time slot can be far from the optimal policy. This is because the locations of the UAV chosen in previous iterations can affect the choices of the UAV's locations in the following iterations. Moreover, as the UAV has a limited energy supply, the greedy search may quickly deplete the energy since it always chooses the "best" next location for the current time slot among those locations reachable with the residual flight energy.

Algorithm 1: OSLR-based UAV's Trajectory Planning

Input: N, $\mathcal{V}$, $\boldsymbol{\Sigma}$, $V_{\max}$, $\boldsymbol{x}_0 = \{\boldsymbol{u}_0, \boldsymbol{K}_0, E_{\max}\}$
Output: $\boldsymbol{K}_M^*$
Initialize $i = 0$
while $i < N - 1$ **do**
  find $S_i(\boldsymbol{x}_i)$
  if $S_i(\boldsymbol{x}_i) \neq \emptyset$ **then**
    $\boldsymbol{s}_i^* = \underset{\boldsymbol{s}_i \in S_i(\boldsymbol{x}_i)}{\operatorname{argmax}}\, J_i^{\mathrm{R}}(\boldsymbol{x}_i)$
    $i = i + 1$, and update $\boldsymbol{x}_i$
  else
    $M = i$, $\boldsymbol{K}_M^* = \boldsymbol{K}_i$ and stop iteration
  end
end

Unlike greedy search, which disregards $J_{i+1}^*(f(\boldsymbol{x}_i, \boldsymbol{s}_i))$, the OSLR approximates $J_{i+1}^*(f(\boldsymbol{x}_i, \boldsymbol{s}_i))$ in (12) as $J_{i+1}^{\mathrm{R}}(f(\boldsymbol{x}_i, \pi_{\mathrm{h}}(\boldsymbol{x}_i))) = \sum_{j=i+1}^{M-1} g(\boldsymbol{x}_j, \pi_h(\boldsymbol{x}_j))$ using the heuristic policy $\pi_{\mathrm{h}}(\cdot)$. The action $\boldsymbol{s}_i$ for a given state $\boldsymbol{x}_i$ is therefore determined by solving the approximate Bellman equation given below

$$J_i^{\mathrm{R}}(\boldsymbol{x}_i) = \max_{\boldsymbol{s}_i \in S_i(\boldsymbol{x}_i)} [g(\boldsymbol{x}_i, \boldsymbol{s}_i) + J_{i+1}^{\mathrm{R}}(f(\boldsymbol{x}_i, \pi_{\mathrm{h}}(\boldsymbol{x}_i)))], \tag{14}$$

where $J_i^{\mathrm{R}}(\boldsymbol{x}_i)$ is the maximum sum reward if actions $\{\boldsymbol{s}_i, \pi_{\mathrm{h}}(\boldsymbol{x}_{i+1}), \cdots, \pi_{\mathrm{h}}(\boldsymbol{x}_M)\}$ are taken starting from state $\boldsymbol{x}_i$. The heuristic policy $\pi_{\mathrm{h}}(\cdot)$ utilizing greedy search is employed for future actions, which simplifies computational requirements. In contrast to the greedy policy $\pi_{\mathrm{h}}(\cdot)$ employed in each iteration, the OSLR algorithm takes into account not only the immediate reward associated with the current action $\boldsymbol{s}_i$, but also considers the long-term consequences and potential rewards resulting from this action $\boldsymbol{s}_i$. Compared to solving the subproblem in (12), the computational complexity of solving the subproblem in (14) is significantly reduced since the policy $\pi_j(\cdot)$ remains fixed as $\pi_h(\cdot)$ for $j \geq i+1$. Hence, the OSLR-based algorithm offers an improved solution compared to greedy search, while maintaining a lower computational complexity than the optimal solution.

Algorithm 1 summarizes the proposed UAV trajectory planning algorithm based on OSLR. The trajectory planning process begins with the given initial state $\boldsymbol{x}_0$, and in each time slot, the state is updated based on the action $\boldsymbol{s}_i^*$ that maximizes the reward $J_i^{\mathrm{R}}(\boldsymbol{x}_i)$. The UAV's states are iteratively explored until the action space $S_M(\boldsymbol{x}_M)$ becomes empty, indicating that the UAV cannot take any further action at state $\boldsymbol{x}_M$ without surpassing the maximum energy limit $E_{\max}$.

V. Simulation Results

In this section, we evaluate the performance of the proposed OSLR-based trajectory planning scheme through simulations. We consider a $40\ \mathrm{m} \times 40\ \mathrm{m}$ square area and equally divide it into 400 small squares of $2\ \mathrm{m} \times 2\ \mathrm{m}$ in size, resulting in

TABLE I
PARAMETER SETTINGS

Parameter	Value
Duration of time slot	$\Delta = 0.08$ s
Maximum flight speed	$V_{\max} = 36$ m/s
Spatial lag parameter in SAR	$\rho = 0.9$
Coefficients of independent variables $\boldsymbol{A}$	$\boldsymbol{\beta} = \mathbf{0}$
Parameter of spatial weight matrix	$\theta = 0.98$
Variance of noise ϵ in SAR	$\sigma^2 = 1.5$
Parameters of the energy consumption model	$\alpha_1 = 0.01$, $\alpha_2 = 1$

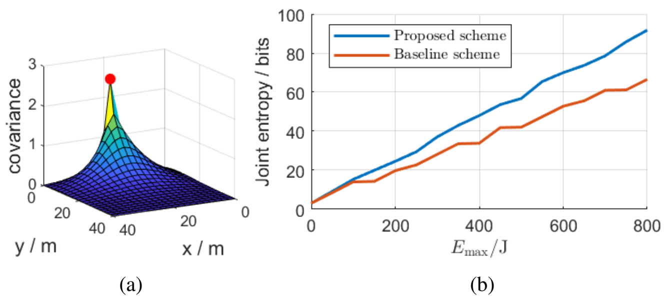


Fig. 2. (a) Values of covariance between Z_{10} and $\boldsymbol{Z}$. (b) Joint entropy versus the maximum energy $E_{\max}$ of proposed and baseline scheme.

$N = 441$ generated grid nodes denoted as $\mathcal{V} = \{\mathbf{v}_1, ..., \mathbf{v}_N\}$. Unless stated otherwise, we set the simulation parameters according to Table I. For comparison, we also evaluate the performance of the greedy search as a baseline scheme.

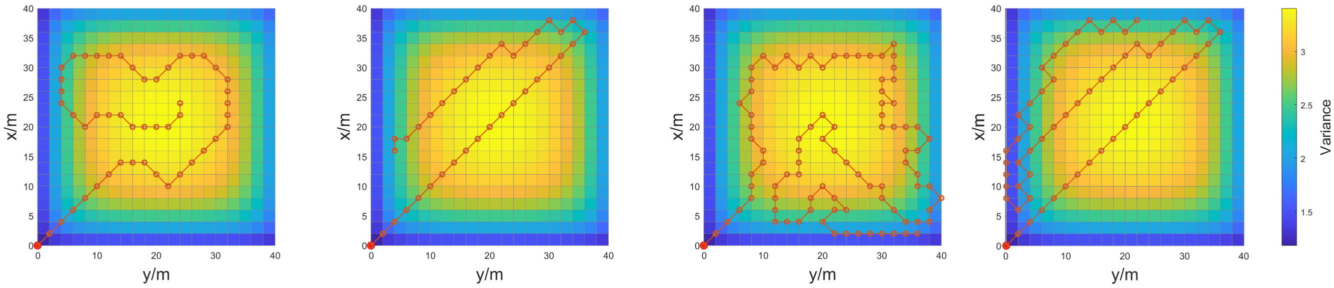
In the simulations, we first consider a covariance matrix generated based on Equation (3) and the parameters in Table I. Figure 2(a) shows the covariance between the pair of random variables Z_{10} and $\boldsymbol{Z}$, which represent the sensed data at $\mathbf{v}_{10} = [18, 0]$ and $\mathcal{V}$, respectively. Note that in Figure 2(a), the peak value, which is highlighted by a red dot, represents the variance of Z_{10}. We observe from Figure 2(a) that the covariance values between Z_{10} and nearby Z_n, particularly around $\mathbf{v}_{10}$, are higher, indicating that the correlation between data sensed in close proximity is higher. Using the simulated covariance matrices, Figure 2(b) further evaluates the joint entropy as a function of the maximum energy $E_{\max}$ of the UAV for both, the proposed and baseline scheme. Notably, the joint entropy increases with the maximum energy $E_{\max}$ for both schemes. This is because, as $E_{\max}$ increases, the UAV can exploit the additional flight energy to cover more area. Moreover, the proposed scheme consistently outperforms the baseline greedy search scheme and the performance gap widens as $E_{\max}$ increases. At $E_{\max} = 800$ J, the proposed scheme achieves an increase of 38% in joint entropy compared to the baseline scheme.

For insights into the performance gains, Figure 3 compares the trajectories obtained by both schemes for $E_{\max} = 1200$ J in (a)-(b) and $E_{\max} = 2000$ J in (c)-(d), where the starting point is $\boldsymbol{u}_0 = [0, 0]$. The heatmap in each subfigure shows the variance of $\boldsymbol{Z}$ for all grid nodes, which peaks in the center and reduces with the distance to the center in all directions. Note that sensing at a position with high variance and hence, large uncertainty, can increase the information in the sensed data. Hence, for all considered setups in Figure 3(a)-(d), the UAV initially flies to the center of the target area. However, the UAV's trajectories differ significantly. Particularly, when $E_{\max} = 1200$ J, Figure 3(a)-(b) show a spiral and a directional flight trajectory for the proposed and the baseline schemes, respectively. This is because the baseline scheme aims at maximizing the information at each position. To this end, the UAV would gather data at spatially distant positions to reduce the correlation among the collected data. Consequently, the UAV flies back and forth on the diagonal of the area at the highest speed since the baseline scheme ignores the impact of each selected position on the future data acquisition, which may rapidly deplete the energy and further limit the total amount of sensed data. Unlike the baseline scheme, the proposed scheme considers the impact of each UAV movement on the subsequent data collection. Consequently, the UAV adopts a reduced speed and a spiral flight pattern to maintain large distances between the data collection points in order to reduce the correlation in the sensed data.

On the other hand, when the maximal allowed energy is increased to $E_{\max} = 2000$ J, we observe from Figure 3(c)-(d) that both, the proposed and the baseline scheme, still show similar flight patterns as in Figure 3(a)-(b). Note that the UAV can utilize the additional energy to explore and sense more positions for both considered schemes. Interestingly, the positions visited by the UAV are well spread over the target area for the proposed scheme. However, the trajectory derived from the baseline scheme is confined in the left upper triangle, which reveals the superior capability of the proposed scheme in trajectory planning. Therefore, the proposed trajectory planning scheme is more advantageous in sensing information maximization than the baseline scheme.

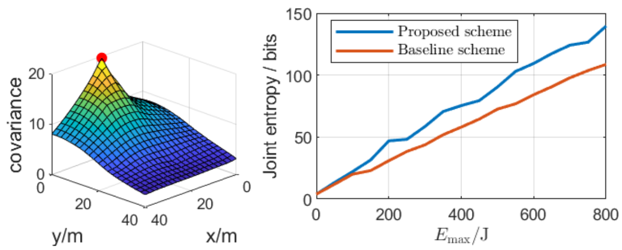
To further validate our results, we employ another covariance matrix calculated by applying real world data to the method from [15], where the daily average wind speeds collected from five stations ('RPT','VAL','DUB','BEL', and 'MAL') in the Republic of Ireland from 1961 to 1978 are used. For the UAV's trajectory planning, the five stations are placed on the $40\ \mathrm{m} \times 40\ \mathrm{m}$ target area, aligned proportionally to their actual positions. Figure 4(a) shows the values of covariance between Z_{10} and $\boldsymbol{Z}$. Figure 4(b) shows the joint entropy versus $E_{\max}$ for the considers schemes. Compared to the baseline scheme, the increase of joint entropy is approximately 28% at $E_{\max} = 800$ J, which confirms the applicability of our proposed scheme.

Figure 5 (a) and (b) show the trajectories obtained by the proposed and baseline scheme with $E_{\max} = 650$ J and starting point $\boldsymbol{u}_0 = [22, 4]$. The trajectory derived from the proposed scheme prioritizes the area with more information in the lower right corner since the proposed scheme considers the impact of each step on future data collection. Furthermore, the greedy search has a tendency to choose a higher speed, which can lead to rapid energy depletion and reduced collected data. To exclude the effect of less time on the amount of the obtained information, Figure 6 shows the joint entropy as a function of the time slots M, where we can see that the proposed scheme still outperforms the baseline scheme, with an increase of about 22% at $M = 45$.



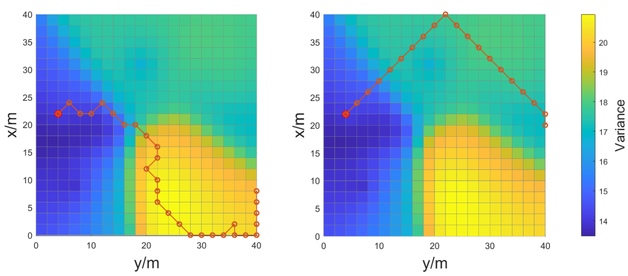
(a) (b) (c) (d)

Fig. 3. UAV's trajectory of (a) the proposed scheme and (b) the baseline scheme, respectively, with $E_{\max} = 1200$ J and of (c) the proposed scheme and (d) the baseline scheme, respectively, with $E_{\max} = 2000$ J.



(a) (b)

Fig. 4. Values of covariance between Z_{10} and $\boldsymbol{Z}$ generated by using the method in [15] based on real-world wind data. (b) Joint entropy versus the maximum energy $E_{\max}$ of proposed and baseline scheme.



(a) (b)

Fig. 5. UAV's trajectory (a) of the proposed scheme, (b) of baseline scheme, with $E_{\max} = 650$ J.

VI. Conclusion

In this paper, we investigated the information maximization problem in the energy-constrained UAV-aided sensed data collection, considering data correlation. The information within the correlated sensed data is quantified using joint entropy. To achieve this objective, we formulated a mixed-integer nonconvex trajectory optimization problem to maximize the joint entropy. To solve the mixed-integer nonconvex problem, we first reformulated the problem into an equivalent DP problem. However, solving the DP problem can be computationally intensive. To this end, we proposed a low-complexity trajectory planning scheme using OSLR, which provides a suboptimal solution for the DP problem. In the simulation results, we compared the performance of the proposed trajectory planning scheme with the baseline scheme based on greedy search. The proposed scheme demonstrates significant improvement, improving the joint entropy by up to 38% compared to the baseline scheme. Furthermore, we applied the proposed scheme on the covariance matrix derived from the real-world data using the method in [15]. The simulation results show that

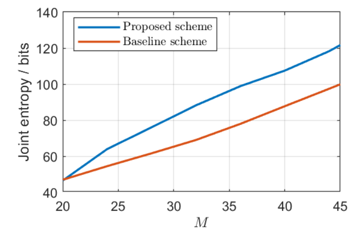


Fig. 6. Joint entropy versus number of time slots M.

our proposed scheme has a performance gain of about 28%, which confirms the applicability of the proposed scheme.

References

[1] A. Matese, P. Toscano, et al., "Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture," *Remote Sensing*, vol. 7, no. 3, pp. 2971-2990, 2015.

[2] T. Guimarães, M. Veronez, et al., "An alternative method of spatial autocorrelation for chlorophyll detection in water bodies using remote sensing," *Sustainability*, vol. 9, no. 3, pp. 416, 2017.

[3] A. Krause, A. Singh, et al., "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *J. Mach. Learn. Res.*, vol. 9, no. 2, 2008.

[4] M. Papadopoulou, B. Raphael, et al., "Hierarchical sensor placement using joint entropy and the effect of modeling error," *Entropy*, vol. 16, no. 9, pp. 5078-5101, 2014.

[5] B. Zhu, et al., "UAV trajectory planning in wireless sensor networks for energy consumption minimization by deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 9540-9554, 2021.

[6] Y. Wang, M. Chen, et al., "Joint optimization of UAV trajectory and sensor uploading powers for UAV-assisted data collection in wireless sensor networks," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 11214-11226, 2021.

[7] M. Nemati, S. Pokhrel, et al., "Modelling Data Aided Sensing With UAVs for Efficient Data Collection," in *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1959-1963, Sept. 2021.

[8] K. Scott, R. Dai, M. Kumar, "Occlusion-aware coverage for efficient visual sensing in unmanned aerial vehicle networks[C]," *IEEE GLOBECOM*, pp. 1-6, 2016.

[9] Y. Liang, et al., "Nonredundant Information Collection in Rescue Applications via an Energy-Constrained UAV," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2945-2958, April 2019.

[10] J. LeSage, "Bayesian estimation of limited dependent variable spatial autoregressive models[J]," *Geographical Analysis*, vol. 32, no. 4, pp. 19-35, 2000.

[11] H. Li, C. Calder, et al., "Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model," *Geographical analysis*, vol. 39, no. 4, pp. 397-375, 2007.

[12] D. Bertsekas, "Dynamic programming and optimal control: Volume I[M]," *Athena scientific*, 2012.

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

[14] D. Silver, et al., "Mastering the game of Go with deep neural networks and tree search[J]," *nature*, vol. 529, no. 7587, pp. 484-489, 2016.

[15] D. Nott, W. Dunsmuir. "Estimation of nonstationary spatial covariance structure[J]," *Biometrika*, vol. 89, no. 4, pp. 819-829, 2002.