Supplementary Material for "LLM Meets the Sky: Heuristic Multi-Agent Reinforcement Learning for Secure Heterogeneous UAV Networks"

Lijie Zheng, Ji He, *Member, IEEE*, Shih Yu Chang, *Senior Member, IEEE*, Yulong Shen, *Member, IEEE*, Dusit Niyato, *Fellow, IEEE*

This document serves as the supplementary material for the paper "LLM Meets the Sky: Heuristic Multi-Agent Reinforcement Learning for Secure Heterogeneous UAV Networks." It presents the detailed mathematical derivation of the SCA–based trajectory optimization baseline, namely the SCA-S2DC algorithm, which is employed in the comparative experiments. These derivations are provided to enhance the transparency and reproducibility of the SCA-S2DC.

I. SCA-based Trajectory Optimization

The original problem model of the Paper is as follows:

$$
\begin{aligned}
\textbf{P1:}\quad & \max_{\boldsymbol{\omega},\boldsymbol{v},\mathbf{P}} \quad F \triangleq \{f_1, -f_2\}, && \text{(1a)}\\
\text{s.t.}\quad & u_k(t) \in [0, D]^2, \quad \forall k \in \mathcal{K}, t \in \mathcal{T}, && \text{(1b)}\\
& \omega_k(t) \in [0, 2\pi), \quad \forall k \in \mathcal{K}, t \in \mathcal{T}, && \text{(1c)}\\
& v_k(t) \le v_{\max}, \quad \forall k \in \mathcal{K}, t \in \mathcal{T}, && \text{(1d)}\\
& \mathrm{Tr}(\mathbf{P}_k(t)\mathbf{P}_k^H(t)) \le P_{\max}, \forall k \in \mathcal{K}, t \in \mathcal{T}, && \text{(1e)}\\
& R_i^c(t) \ge R_{e,i}^c(t), \forall k \in \mathcal{K}, e \in \mathcal{E}, t \in \mathcal{T}, && \text{(1f)}\\
& d_{k,k'}(t) \ge d_c, k, k' \in \mathcal{K}, k \ne k', t \in \mathcal{T}, && \text{(1g)}\\
& \sum_{i \in \mathcal{I}} \mathbf{S}_{k,i}^{\mathcal{I}}(t) \le N_k^s, \forall k \in \mathcal{K}, t \in \mathcal{T}, && \text{(1h)}\\
& \sum_{k \in \mathcal{K}} \mathbf{S}_{k,i}^{\mathcal{I}}(t) \le 1, \forall i \in \mathcal{I}, t \in \mathcal{T}. && \text{(1i)}
\end{aligned}
$$

To render the problem computationally tractable, we first reformulate the original multi-objective optimization problem into a single-objective problem by introducing appropriate weighting coefficients, thereby enabling an explicit trade-off between different objectives. Subsequently, leveraging the block coordinate descent (BCD) framework, we decompose the resulting problem into two tightly coupled subproblems: secrecy precoding design and UAV trajectory optimization [1]. These two subproblems are solved in an alternating iterative manner. Specifically, we first fix the UAV trajectories and solve the precoding subproblem using the S2DC algorithm. Then, eigenvalue decomposition is performed on the obtained outer product matrix to recover the rank-1 approximation, thereby obtaining the secrecy precoding vector. Next, with the precoding vector held fixed, we optimize the UAV trajectories by applying the successive convex approximation (SCA) method to handle the non-convex constraints and objective terms.

A. Problem Reformulation

We employ a linear weighting method [2], [3] to transform the multi-objective optimization problem into a single-objective optimization problem because this method has lower algorithmic complexity and can quickly produce effective results. Therefore, the multi-objective optimization problem is reconstructed as

$$\rho(\boldsymbol{\omega}, \boldsymbol{v}, \mathbf{P}) \triangleq w^{\mathrm{sr}} f_1 - w^{\mathrm{ec}} f_2, \tag{2}$$

where the non-negative parameters w^{sr} and w^{ec} represent the weighting factors for secrecy rate and propulsion energy consumption, respectively. Note that the weight coefficients need be adjusted based on the numerical range between objectives and the user preferences on the objectives.

B. Secrecy Precoding

Given the UAV trajectories $\hat{\boldsymbol{\omega}}$, $\hat{\boldsymbol{v}}$, the secrecy precoding optimization problem can be expressed as:

$$
\begin{aligned}
\textbf{P2:}\quad & \max_{\mathbf{P}} \quad \rho(\hat{\boldsymbol{\omega}}, \hat{\boldsymbol{v}}, \mathbf{P}) \triangleq w^{\mathrm{sr}} f_1, && \text{(3a)}\\
\text{s.t.}\quad & \text{(1e)}, \text{(1f)}.
\end{aligned}
$$

This problem can be solved efficiently using the S2DC algorithm designed in the paper. By using eigenvalue decomposition, we can recover the rank-1 solution from the resulting outer product matrix and obtain the precoded vector.

C. UAV Trajectory Control

Given the UAV precoding $\hat{\mathbf{P}}$, the UAV trajectory optimization problem can be expressed as:

$$
\begin{aligned}
\textbf{P3:}\quad & \max_{\boldsymbol{\omega},\boldsymbol{v}} \quad \rho(\boldsymbol{\omega}, \boldsymbol{v}, \hat{\mathbf{P}}) \triangleq w^{\mathrm{sr}} f_1 - w^{\mathrm{ec}} f_2 && \text{(4a)}\\
\text{s.t.}\quad & \text{(1b)} - \text{(1d)}, \text{(1g)} - \text{(1i)}.
\end{aligned}
$$

To make the current problem easier to handle, we need to perform some equivalent transformations. First, the existence of f_1 makes (4) a max-min problem. Without loss of generality, we introduce slack variable φ to address this in the optimization problem. Second, to make the optimization variables easier to handle, we use the UAV's position to represent its trajectory. Finally, since (1h) and (1i) are deterministic rules based on channel quality, they can be ignored without affecting convexity (the specific implementation method has been added below Eq. (9) in the revised paper).

Therefore, the optimization problem can ultimately be transformed into the following form:

$$\textbf{P4:} \quad \max_{\boldsymbol{u}} \quad \rho(\boldsymbol{u}, \hat{\mathbf{P}}) \triangleq w^{\text{sr}}\varphi - w^{\text{ec}} f_2 \tag{5a}$$
$$\text{s.t.} \quad \alpha_{k,i} R_k^{\text{sr,c}}(t) + R_{k,i}^{\text{sr,p}}(t) \geq \varphi, \forall k \in \mathcal{K}, i \in \mathcal{I}, t \in \mathcal{T}, \tag{5b}$$
$$\|u_k(t+1) - u_k(t)\|_2 \leq v_{\max}\Delta t, \tag{5c}$$
$$(1\text{b}), (1\text{g}).$$

where (5c) is the velocity constraint after the equivalent transformation.

Since f_2 in the objective function is non-convex, and constraints (5b) and (1g) are also non-convex, sub-problem **P4** is therefore non-convex. To solve this non-convex sub-problem, we adopt the SCA method to optimize the UAV trajectory.

1) Handling the non-convexity of f_2: In objective f2, the non-convexity stems from the energy consumption model, which is shown below:

$$P_k(v_k(t)) = \frac{1}{2} d_0 \rho_a s_{\text{sol}} A v_k(t)^3 + P_0 \left(1 + \frac{3 v_k(t)^2}{v_{\text{tip}}^2}\right) + P_1 \left(\sqrt{1 + \frac{v_k(t)^4}{4 v_0^4}} - \frac{v_k(t)^2}{2 v_0^2}\right)^{\frac{1}{2}}. \tag{6}$$

As we can see from the Eq. (6), the first and second terms are convex, while the third term is non-convex. To address this non-convex term, we introduce slack variables $\{y_k \geq 0\}$ such that [4]

$$P_1 \left(\sqrt{1 + \frac{v_k^4(t)}{4 v_0^4}} - \frac{v_k^2(t)}{2 v_0^2}\right)^{\frac{1}{2}} \leq y_k(t), \forall k \in \mathcal{K}, t \in \mathcal{T}. \tag{7}$$

Therefore, the third term in Eq. (6) can be replaced by the linear term y_k, with the addition of constraint (7).

For the newly added constraints, we perform a continuous square transformation on the inequality and obtain the following equivalent representation:

$$1 + \frac{v_k^4(t)}{4 v_0^4} \leq \left(\frac{y_k^2(t)}{P_1^2} + \frac{v_k^2(t)}{2 v_0^2}\right)^2. \tag{8}$$

Then, we expand and rearrange the RHS of the inequality to obtain

$$\frac{1}{y_k^2(t)} \leq \frac{y_k^2(t)}{P_1^4} + \frac{v_k^2(t)}{P_1^2 v_0^2}. \tag{9}$$

where both the LHS and RHS sides of the formula are convex functions, but the constraint itself is still non-convex.

By substituting the velocity $v_k(t)$ with the displacement $\|u_k(t) - u_k(t-1)\|_2/\Delta t$, Eq. (9) can be transformed into the following mathematically equivalent form:

$$\frac{1}{y_k^2(t)} \leq \mathcal{G}(y_k(t), u_k(t)), \tag{10}$$

where $\mathcal{G}(y_k(t), u_k(t)) \triangleq \frac{y_k^2(t)}{P_1^4} + \frac{\|u_k(t) - u_k(t-1)\|_2^2}{P_1^2 v_0^2 \Delta t^2}$.

Since the function $\mathcal{G}(\cdot)$ is jointly convex with respect to y_k and $u_k(t)$, we apply SCA by replacing it with its first-order Taylor expansion at the r-th iteration point $(y_k^{(r)}, u_k^{(r)}(t))$:

$$\frac{1}{y_k^2} \leq \mathcal{G}(y_k^{(r)}, u_k^{(r)}(t)) + \nabla_y (y_k - y_k^{(r)}) + \nabla_u^T (u_k - u_k^{(r)}), \tag{11}$$

specifically expanded as (12).

2) Handling the non-convexity of (5b)*:* Substituting the channel $\mathbf{h}_{k,x}(t) = \sqrt{10^{-\frac{1}{10} \times \ell_{k,x}(t)}}\, \hat{\mathbf{h}}_{k,x}(t), x \in \{\mathcal{I}, \mathcal{E}\}$ into Eq. (5b), and letting large scale fading be $g_{k,x} = 10^{-\frac{1}{10} \times \ell_{k,x}}$, then $R_k^{\text{sr,c}}$ can be expressed as

$$R_k^{\text{sr,c}} = \log(1 + \frac{g_{k,i}\ A}{g_{k,i}\ B + C}) - \log(1 + \frac{g_{k,e}\ D}{g_{k,e}\ E + F}), \tag{13}$$

where $A = \mathbf{S}_{k,i}^{\mathcal{I}} \left|\hat{\mathbf{h}}_{k,i}^H \mathbf{p}_k^c\right|^2$, $B = \mathbf{S}_{k,i}^{\mathcal{I}} \sum_{i' \in \mathcal{I}_k} \left|\hat{\mathbf{h}}_{k,i}^H \mathbf{p}_{k,i'}^p\right|^2$, $C = \sum_{k' \in \mathcal{K} \setminus \{k\}} \mathbf{A}_{k',i}^{\mathcal{I}}\ g_{k',i} \left|\hat{\mathbf{h}}_{k',i}^H \mathbf{P}_{k'}\right|^2 + \sigma_i^2$, $D = \mathbf{A}_{k,e}^{\mathcal{E}} \left|\hat{\mathbf{h}}_{k,e}^H \mathbf{p}_k^c\right|^2$, $E = \mathbf{A}_{k,e}^{\mathcal{E}} \sum_{i' \in \mathcal{I}_k} \left|\hat{\mathbf{h}}_{k,e}^H \mathbf{p}_{k,i'}^p\right|^2$, and $F = \mathbf{A}_{k,e}^{\mathcal{E}} \sum_{i' \in \mathcal{I}_k} \left|\hat{\mathbf{h}}_{k,e}^H \mathbf{p}_{k,i'}^p\right|^2$.

Similarly, $R_{k,i}^{\text{sr,p}}$ can be represented as

$$R_{k,i}^{\text{sr,p}} = \log(1 + \frac{g_{k,i}\ G}{g_{k,i}\ H + C}) - \log(1 + \frac{g_{k,e}\ I}{g_{k,e}\ J + F}) \tag{14}$$

where $G = \mathbf{S}_{k,i}^{\mathcal{I}} \left|\hat{\mathbf{h}}_{k,i}^H \mathbf{p}_{k,i}^p\right|^2$, $H = \mathbf{S}_{k,i}^{\mathcal{I}} \sum_{i' \in \mathcal{I}_k \setminus \{i\}} \left|\hat{\mathbf{h}}_{k,i}^H \mathbf{p}_{k,i'}^p\right|^2$, $I = \mathbf{A}_{k,e}^{\mathcal{E}} \left|\hat{\mathbf{h}}_{k,e}^H \mathbf{p}_{k,i}^p\right|^2$, and $J = \mathbf{A}_{k,e}^{\mathcal{E}} (\left|\hat{\mathbf{h}}_{k,e}^H \mathbf{p}_k^c\right|^2 + \sum_{i' \in \mathcal{I}_k \setminus \{i\}} \left|\hat{\mathbf{h}}_{k,e}^H \mathbf{p}_{k,i'}^p\right|^2)$.

We further perform a common denominator operation on $R_k^{\text{sr,c}}$ to obtain

$$R_k^{\text{sr,c}} = \log(g_{k,i}\ (A + B) + C) - \log(g_{k,i}\ B + C) - \tag{16}$$
$$\log(g_{k,e}\ (D + E) + F) + \log(g_{k,e}\ E + F). \tag{17}$$

Then, by differentiating $R_k^{\text{sr,c}}$, we obtain

$$\nabla_{u_k} R_k^{\text{sr,c}} = \nabla_{u_k} R_i^c - \nabla_{u_k} R_{e,i}^c. \tag{18}$$

According to the chain rule, $\nabla_{u_k} R_i^c$ is derived as (15), where

$$\frac{\partial\, R_i^{\text{c}}}{\partial\, g_{k,i}} = \frac{1}{\ln 2} (\frac{A + B}{g_{k,i}\ (A + B) + C} - \frac{B}{g_{k,i}\ B + C}), \tag{19}$$

$$\frac{\partial\, g_{k,i}}{\partial\, \ell_{k,i}} = -\frac{\ln 10}{10} \cdot 10^{-\frac{1}{10} \times \ell_{k,i}}, \tag{20}$$

$$\frac{\partial\, \ell_{k,i}}{\partial\, P_{k,i}^{\text{LoS}}} = \eta^{\text{LoS}} - \eta^{\text{NLoS}}, \tag{21}$$

$$\frac{\partial\, P_{k,i}^{\text{LoS}}}{\partial\, d_{k,i}} = b\, P_{k,i}^{\text{LoS}}\, (1 - P_{k,i}^{\text{LoS}}) \cdot (-\frac{H_{\text{UAV}}}{d_{k,i}^2 \sqrt{1 - (H_{\text{UAV}}/d_{k,i})^2}}), \tag{22}$$

$$\frac{\partial\, \ell_{k,i}}{\partial\, \text{FL}_{k,i}} = 1, \tag{23}$$

$$\frac{\partial\, \text{FL}_{k,i}}{\partial\, d_{k,i}} = \frac{20}{d_{k,i} \ln 10}, \tag{24}$$

$$\frac{\partial\, d_{k,i}}{\partial u_k} = \frac{\hat{u}_k - \hat{u}_i}{d_{k,i}}. \tag{25}$$

$$\frac{1}{y_k^2} \leq \left(\frac{(y_k^{(r)})^2}{P_1^4} + \frac{\|u_k^{(r)}(t) - u_k^{(r)}(t-1)\|^2}{P_1^2 v_0^2 \Delta t^2} \right) + \frac{2y_k^{(r)}}{P_1^4}(y_k - y_k^{(r)}) + \frac{2(u_k^{(r)}(t) - u_k^{(r)}(t-1))^T}{P_1^2 v_0^2 \Delta t^2}(u_k(t) - u_k(t-1) - (u_k^{(r)}(t) - u_k^{(r)}(t-1))). \tag{12}$$

$$\nabla_{u_k} R_i^c = \frac{\partial R_i^c}{\partial g_{k,i}} \cdot \frac{\partial g_{k,i}}{\partial \ell_{k,i}} \cdot \left(\frac{\partial \ell_{k,i}}{\partial P_{k,i}^{\mathrm{LoS}}} \cdot \frac{\partial P_{k,i}^{\mathrm{LoS}}}{\partial d_{k,i}} + \frac{\partial \ell_{k,i}}{\partial \mathrm{FL}_{k,i}} \cdot \frac{\partial \mathrm{FL}_{k,i}}{\partial d_{k,i}}\right) \cdot \frac{\partial d_{k,i}}{\partial u_k}. \tag{15}$$

$$\beta_i^c \triangleq \frac{\partial R_{k,i}^{\mathrm{sr,c}}}{\partial g_{k,i}} \cdot \frac{\partial g_{k,i}}{\partial \ell_{k,i}} \cdot \left(\frac{\partial \ell_{k,i}}{\partial P_{k,i}^{\mathrm{LoS}}} \cdot \frac{\partial P_{k,i}^{\mathrm{LoS}}}{\partial d_{k,i}} + \frac{\partial \ell_{k,i}}{\partial \mathrm{FL}_{k,i}} \cdot \frac{\partial \mathrm{FL}_{k,i}}{\partial d_{k,i}}\right) \cdot \frac{1}{d_{k,i}}. \tag{26}$$

For convenience, let β_i^c be (26), then $\nabla_{u_k} R_k^{\mathrm{sr,c}}$ can be represented as

$$\nabla_{u_k} R_k^{\mathrm{sr,c}} = \beta_i^c(\hat{u}_k - \hat{u}_i) - \beta_{e,i}^c(\hat{u}_k - \hat{u}_e), \tag{27}$$

where $\hat{u}_k = [x_k, y_k]$, $\hat{u}_i = [x_i, y_i]$, $\hat{u}_e = [x_e, y_e]$ represent the planar coordinates of UAV k, GT i, and Eve e, respectively. Through similar derivation, we can obtain

$$\nabla_{u_k} R_{k,i}^{\mathrm{sr,p}} = \beta_i^p(\hat{u}_k - \hat{u}_i) - \beta_{e,i}^p(\hat{u}_k - \hat{u}_e). \tag{28}$$

In the r-th SCA iteration, the secrecy rates of the common and private streams are approximated by their first-order Taylor expansions w.r.t. the UAV trajectory, i.e.,

$$R_k^{\mathrm{sr,c},(r)}(\hat{u}_k) = R_k^{\mathrm{sr,c}}(\hat{u}_k^{(r)}) + (\nabla_{\hat{u}_k} R_k^{\mathrm{sr,c}}(\hat{u}_k^{(r)}))^T(\hat{u}_k - \hat{u}_k^{(r)}), \tag{29}$$

$$R_{k,i}^{\mathrm{sr,p},(r)}(\hat{u}_k) = R_{k,i}^{\mathrm{sr,p}}(\hat{u}_k^{(r)}) + (\nabla_{\hat{u}_k} R_{k,i}^{\mathrm{sr,p}}(\hat{u}_k^{(r)}))^T(\hat{u}_k - \hat{u}_k^{(r)}). \tag{30}$$

Hence, the secrecy constraint is convexified as

$$\varphi \leq \alpha_{k,i} R_k^{\mathrm{sr,c},(r)}(\hat{u}_k) + R_{k,i}^{\mathrm{sr,p},(r)}(\hat{u}_k), \forall k \in \mathcal{K}, i \in \mathcal{I}, t \in \mathcal{T}. \tag{31}$$

3) Handling the non-convexity of (1g)*:* This constraint can be equivalently written as

$$-\|u_k - u_{k'}\|_2^2 \leq -d_c^2, \forall k \neq k', t \in \mathcal{T}. \tag{32}$$

Define $z_{k,k'} = u_k - u_{k'}$ and $f(z) = -\|z\|_2^2$, which is a concave function. In the r-th SCA iteration, given $u_k^{(r)}$ and $u_{k'}^{(r)}$, we have $z^{(r)} = u_k^{(r)} - u_{k'}^{(r)}$. The first-order Taylor expansion of $f(z)$ at $z_{kk'}^{(r)}$ yields a global affine upper bound:

$$f(z_{kk'}) \leq -\|z_{kk'}^{(r)}\|_2^2 - 2(z_{kk'}^{(r)})^T(z_{kk'} - z_{kk'}^{(r)}). \tag{33}$$

Substituting $z_{kk'} = u_k - u_{k'}$ gives the following convex approximation of the collision avoidance constraint:

$$-\|z_{kk'}^{(r)}\|_2^2 - 2(z_{kk'}^{(r)})^T[(u_k - u_k^{(r)}) - (u_{k'} - u_{k'}^{(r)})] \leq -d_c^2. \tag{34}$$

By replacing the non-convex terms in f_2 and non-convex constraints (1g), (5b) with the corresponding lower bounds obtained in the r-th iteration, we obtain the following optimization problem:

$$\textbf{P5:} \quad \max_{\boldsymbol{u}} \quad \rho(\boldsymbol{u}, \hat{\mathbf{P}}) \triangleq w^{\mathrm{sr}}\varphi - w^{\mathrm{ec}} f_2 \tag{35a}$$

$$\text{s.t.} \quad (1b), (5c), (11), (31), (34).$$

Now, this problem is a convex optimization problem, which can be solved efficiently using CVX. Since the UAV trajectory actions in our original formulation are discretized, for a fair comparison we post-process the continuous trajectories obtained by SCA by projecting them back onto the discrete action space used in our framework.

Algorithm 1: Joint Trajectory and Precoding Optimization via SCA-S2DC

1: **Input:** Coordinates of UAV, GT, and Eve; channel; weighting factors $w^{\mathrm{sr}}, w^{\mathrm{ec}}$.
2: **Output:** Trajectory $\{\boldsymbol{u}_k^*\}$ and precoding matrices $\mathbf{P}^*$.
3: **Initialization:**
4: Initialize a feasible trajectory $\{\boldsymbol{u}_k^{(0)}\}$ for all $k \in \mathcal{K}$ and $t \in \mathcal{T}$;
5: Set outer iteration index $l = 0$, maximum iterations $L_{\max}$, tolerance $\epsilon > 0$;
6: Compute the initial precoding matrices $\mathbf{P}^{(0)}$ by running the S2DC algorithm with the channels induced by $\{\boldsymbol{u}_k^{(0)}\}$;
7: **repeat**
8: $\quad l \leftarrow l + 1$;
9: $\quad$ Given the current trajectory $\{\boldsymbol{u}_k^{(l-1)}\}$, the optimized outer product matrix $\{\mathbf{P}^{c,(l)}, \mathbf{P}^{p,(l)}\}$ is obtained through S2DC.
10: $\quad$ Perform eigenvalue decomposition and recover rank-one precoding vectors $\{\mathbf{p}_k^{c,(l)}\}$, $\{\mathbf{p}_{k,i}^{p,(l)}\}$;
11: $\quad$ Given $\{\mathbf{p}_k^{c,(l)}\}$, $\{\mathbf{p}_{k,i}^{p,(l)}\}$, solve (35) to obtain the UAV trajectory $\{\boldsymbol{u}_k^{(l)}\}$.
12: **until** $|\rho^{(l)} - \rho^{(l-1)}| \leq \epsilon$ or $l \geq L_{\max}$;
13: **Return:** $\{\boldsymbol{u}_k^*, \mathbf{p}_k^c, \mathbf{p}_{k,i}^p\}$.

The complete SCA-S2DC algorithm flow is summarized as Algorithm 1.

References

[1] G. Sun *et al.*, "Multi-objective optimization for multi-uav-assisted mobile edge computing," *IEEE Transactions on Mobile Computing*, vol. 23, no. 12, pp. 14 803–14 820, 2024.

[2] I. P. Stanimirovic, M. L. Zlatanovic, and M. D. Petkovic, "On the linear weighted sum method for multi-objective optimization," *Facta Acta Univ*, vol. 26, no. 4, pp. 49–63, 2011.

[3] I. Alsyouf and S. Hamdan, "A multi-objective optimization of maintenance policies using weighted comprehensive criterion method (wccm)," in *2017 7th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO)*. IEEE, 2017, pp. 1–4.

[4] Y. Zeng, J. Xu, and R. Zhang, "Energy minimization for wireless communication with rotary-wing uav," *IEEE transactions on wireless communications*, vol. 18, no. 4, pp. 2329–2345, 2019.