$$\frac{1}{\nu P_{\max}}\ell_\nu^*(\mathbf{s};\hat{r}_U,1) {=} \min_{\Delta,\mathbf{p}_U,t_p} -\ln(eT)\Delta {+} \frac{1}{P_{\max}}\int_0^\Delta P_{\text{mob}}\Big(\sqrt{r_U'(\eta)^2 {+} r_U^2(\eta)\theta_U'(\eta)^2}\Big)\mathrm{d}\eta \text{ s.t.} constraints \tag{1}$$

where $T=\exp\{-1-(\frac{1}{\nu}-P_{\text{avg}})/P_{\max}\}$. T varies between $T=0$ (when $\nu\to 0$) to $T=1$ (when $\nu\to\infty$ and $P_{\text{avg}}\to P_{\max}$. Let $\Delta_T(\mathbf{s};\hat{r}_U)$ and $E_T(\mathbf{s};\hat{r}_U)$ the optimal delay and energy cost for a certain T (and associated trajectory)/

- Radii levels: K_R+1 levels $r_j=aj/K_R$ for $j=0,\dots,K_R$ (encoding UAV positions)

- Radii levels: G_R+1 levels $r_\ell=a\ell/G_R$ for $\ell=0,\dots,G_R$ (encoding GN radii positions)

- Angular levels (for GN request generation): for radius level ℓ, $G_{A,\ell}+1$ angular levels $\phi_{\ell,z}=z/G_{A,\ell}\pi$, $z=0,\dots,G_{A,\ell}$ (note that value function is symmetric (with symmetric trajectries), wrt ϕ hence only $\phi\in[0,\pi]$ is considered).

- Radial velocity level: $2K_V+1$ levels $v_u=u/K_VV_{\max}$ where $u=-K_V,\dots,K_V$.

Trajectory optimziation during comm:

Step 1: For every $(j,\ell,z,\hat{j})$ (encoding the current UAV radius r_j, GN position $(r_\ell,\phi_{\ell,z})$, next UAV radius position $r_{\hat{j}}$) find a set of $Q+1$ trajectories by solving the problem

$$\min_{\Delta,\mathbf{p}_U,t_p} -\ln(eT_q)\Delta {+} \frac{1}{P_{\max}}\int_0^\Delta P_{\text{mob}}\Big(\sqrt{r_U'(\eta)^2 {+} r_U^2(\eta)\theta_U'(\eta)^2}\Big)\mathrm{d}\eta \text{ s.t.} constraints \tag{2}$$

via HCSO. Here $T_q=q/Q$ for $q=0,\dots Q$. $T_q=0$ ($\nu=0$) is a special case:

$$\min_{\Delta,\mathbf{p}_U,t_p}\Delta \text{ s.t.} constraints \tag{3}$$

Note that this involves $Q+1$ HCSO calls, for every state-outer action $(j,\ell,z,\hat{j})$. Since there are $(K_R+1)^2\sum_{j=0}^{G_R}G_{A,j}$ state - outer actions, total number of HCSO calls is

$$(Q+1)(K_R+1)^2\sum_{j=0}^{G_R}G_{A,j}$$

For each trajectory indexed by $q=0\dots Q$, you should have its delay and energy costs:

$$\Delta_q(j,\ell,z,\hat{j}),\ \ E_q(j,\ell,z,\hat{j})$$

Let $V_C(j)$ be the value function when a request arrives with the UAV in radial position

$r_U = aj/K_R$, averaged out with respect to the request position. This is computed as

$$V_C(j) = \int_0^{2\pi} \frac{1}{2\pi} \int_0^a \frac{2r'}{a^2} V(r_U, r', \psi') \mathrm{d}r' \mathrm{d}\psi'$$
$$= \int_0^{\pi} \frac{1}{\pi} \int_0^a \frac{2r'}{a^2} V(r_U, r', \psi') \mathrm{d}r' \mathrm{d}\psi'$$

(angular symmetry) and

$$\int_0^a \frac{2r'}{a^2} V(r_U, r', \psi') \mathrm{d}r' = \sum_{\ell=0}^{G_R-1} \int_{a\ell/G_R}^{a(\ell+1)/G_R} \frac{2r'}{a^2} V(r_U, r', \psi') \mathrm{d}r'$$

and approximating $V(r_U, r', \psi')$ for $r' \in [a\ell/G_R, a(\ell+1)/G_R]$ via linear interpolation as

$$V(r_U, r', \psi') \approx [(\ell+1) - r'G_R] V(r_U, a\ell/G_R, \psi') + [r'G_R - \ell] V(r_U, a(\ell+1)/G_R, \psi')$$

we obtain

$$\int_0^a \frac{2r'}{a^2} V(r_U, r', \psi') \mathrm{d}r' = \sum_{\ell=0}^{G_R} p_{R,\ell} V(r_U, a\ell/G_R, \psi')$$

where $p_{R,\ell}$ is defined as

$$p_{R,0} = \frac{1/3}{G_R^2},\ p_{R,\ell} = \frac{2\ell}{G_R^2}\ \forall \ell = 1 \dots G_R - 1,\ p_{R,G_R} = \frac{G_R - 1/3}{G_R^2}$$

(note that they sum to one) Therefore

$$V_C(j) = \int_0^{2\pi} \frac{1}{2\pi} \int_0^a \frac{2r'}{a^2} V(r_U, r', \psi') \mathrm{d}r' \mathrm{d}\psi'$$
$$= \sum_{\ell=0}^{G_R} p_{R,\ell} \int_0^{\pi} \frac{1}{\pi} V(r_U, a\ell/G_R, \psi') \mathrm{d}\psi'$$
$$= \sum_{\ell=0}^{G_R} p_{R,\ell} \sum_{z=0}^{G_{A,\ell}-1} \int_{z\pi/G_{A,\ell}}^{(z+1)\pi/G_{A,\ell}} \frac{1}{\pi} V(r_U, a\ell/G_R, \psi') \mathrm{d}\psi'$$

and approximating $V(r_U, a\ell/G_R, \psi')$ for $\psi' \in [z\pi/G_{A,\ell}, (z+1)\pi/G_{A,\ell}]$ via linear interpolation

$$V(r_U, a\ell/G_R, \psi') \approx [(z+1) - G_{A,\ell}\psi'/\pi] V(r_U, a\ell/G_R, z\pi/G_{A,\ell}) + (G_{A,\ell}\psi'/\pi - z) V(r_U, a\ell/G_R, (z+1)\pi/G_{A,\ell})$$

we obtain

$$\int_{z\pi/G_{A,\ell}}^{(z+1)\pi/G_{A,\ell}} \frac{1}{\pi} V_{C,i}(r_U, a\ell/G_R, \psi') \mathrm{d}\psi' = \frac{V(r_U, a\ell/G_R, z\pi/G_{A,\ell}) + V(r_U, a\ell/G_R, (z+1)\pi/G_{A,\ell})}{2G_{A,\ell}}$$

and

$$\sum_{z=0}^{G_{A,\ell}-1} \int_{z\pi/G_{A,\ell}}^{(z+1)\pi/G_{A,\ell}} \frac{1}{\pi} V_{C,i}(r_U, a\ell/G_R, \psi')\mathrm{d}\psi' = \sum_{z=0}^{G_{A,\ell}} p_{A,\ell,z} V(r_U, a\ell/G_R, z\pi/G_{A,\ell})$$

where $p_{A,\ell,z}$ is defined as

$$p_{A,\ell,0} = \frac{1}{2G_{A,\ell}},\ p_{A,\ell,z} = \frac{1}{G_{A,\ell}}\ \forall z = 1 \dots G_{A,\ell} - 1,\ p_{A,\ell,G_{A,\ell}} = \frac{1}{2G_{A,\ell}}$$

(note that they sum to one for each ℓ) yielding

$$V_C(j) = \sum_{\ell=0}^{G_R} p_{R,\ell} \sum_{z=0}^{G_{A,\ell}} p_{A,\ell,z} V(r_U, a\ell/G_R, z\pi/G_{A,\ell})$$

(note that $\sum_{\ell=0}^{G_R} p_{R,\ell} \sum_{z=0}^{G_{A,\ell}} p_{A,\ell,z} = 1$, as it should be)

Algorithm 1 Value Iteration: $(O^*, U^*, g(\nu), \bar{E}, \bar{T}) = \text{VITER}(\nu)$

Initialization: i=0; waiting state value function $V_{W,i}(j)$=0, total energy cost $E_{W,i}(j)$=0, and total time cost $T_{W,i}(j)$=0, $\forall j = 0, \ldots, K_R$; comm state value function $V_{C,i}(j)$=0, total energy cost $E_{C,i}(j)$=0, and total time cost $T_{C,i}(j)$=0, $\forall j = 0, \ldots, K_R$; stop criterion δ.

Inner optimization in waiting states: $\forall j = 0, \ldots, K_R, \forall u = -K_V, \ldots, K_V$, calculate

$$\ell_\nu^*(j; u) = \nu \left(P_{\text{mob}}(\max\{|u/K_V|V_{\max}, V_{\min}\}) - P_{\text{avg}}\right) \Delta_0;$$

compute energy cost

$$e^*(j; u) = P_{\text{mob}}(\max\{|u/K_V|V_{\max}, V_{\min}\})\Delta_0$$

and time cost

$$t^*(r_U; v_r) = \Delta_0.$$

Inner optimization in communication states: $\forall (j, \ell, a, \hat{j})$, with $j = 0 \ldots K_R$, $\hat{j} = 0 \ldots K_R$, $\ell = 0 \ldots, G_R$, $a = 0 \ldots G_{A,\ell}$: calculate

$$\ell_\nu^*(j, \ell, a, \hat{j}, 1) = \min_{q=0,\ldots,Q} (1 - \nu P_{avg})\Delta_q(j, \ell, a, \hat{j}) + \nu E_q(j, \ell, a, \hat{j});$$

with the minimizer q^*, compute energy/time costs

$$e^*(j, \ell, a, \hat{j}) = E_{q^*}(j, \ell, a, \hat{j}), \;\; t^*(j, \ell, a, \hat{j}) = \Delta_{q^*}(j, \ell, a, \hat{j}).$$

repeat

  for each $j = 0 \ldots, K_R$ **do** ▷ Outer optimization in waiting states

$$V_{W,i+1}(j) \leftarrow \min_{u=-K_V,\ldots,K_V} \left[\ell_\nu^*(j; u) + e^{-\Lambda' \Delta_0}[\alpha V_{W,i}(\tilde{j}) + (1-\alpha)V_{W,i}(\tilde{j}+1)] + (1 - e^{-\Lambda' \Delta_0})[\alpha V_{C,i}(\tilde{j}) + (1-\alpha)V_{C,i}(\tilde{j}+1)]\right],$$

  where for each u, $\alpha \in [0, 1]$ and (nearest index) $\tilde{j}$ are such that $|j/K_R + u/K_V\Delta_0| = \alpha\tilde{j}/K_R + (1-\alpha)(\tilde{j}+1)/K_R$ (linear interpolation).

    $O_{i+1}(j) \leftarrow u^*$, where u^* is the arg min, with α and $\tilde{j}$ (below) corresponding linear interpolation.

    $E_{i+1}(j) \leftarrow e^*(j; u) + e^{-\Lambda' \Delta_0}[\alpha E_{W,i}(\tilde{j}) + (1-\alpha)E_{W,i}(\tilde{j}+1)] + (1 - e^{-\Lambda' \Delta_0})[\alpha E_{C,i}(\tilde{j}) + (1-\alpha)E_{C,i}(\tilde{j}+1)]$.

    $T_{i+1}(j) \leftarrow t^*(j; u) + e^{-\Lambda' \Delta_0}[\alpha T_{W,i}(\tilde{j}) + (1-\alpha)T_{W,i}(\tilde{j}+1)] + (1 - e^{-\Lambda' \Delta_0})[\alpha T_{C,i}(\tilde{j}) + (1-\alpha)T_{C,i}(\tilde{j}+1)]$.

  end for

  for each $j = 0, \ldots, K_R$ **do** ▷ Outer optimization in communication states

    for each $\ell = 0, \ldots, G_R$, $a = 0, \ldots, G_{A,\ell}$ **do**

      $TEMP_V(j, \ell, a) \leftarrow \min\{\frac{L}{R_{GB}(\ell/G_R)} + V_{W,i+1}(j), \min_{\hat{j}=0\ldots,K_R} \left[\ell_\nu^*(j, \ell, a, \hat{j}, 1) + V_{W,i+1}(\hat{j})\right]$

      $U_{i+1}(j, \ell, a) \leftarrow \hat{j}^*$, where $\hat{j}^*$ is the arg min and scheduling decision $\xi_{i+1}(j, \ell, a) \leftarrow \xi^*$.

      $TEMP_{energy}(j, \ell, a) \leftarrow \xi^* \cdot e^*(j, \ell, a, \hat{j}^*) + E_{W,i+1}(\hat{j}^*)$

      $TEMP_{time}(j, \ell, a) \leftarrow \xi^* \cdot t^*(j, \ell, a, \hat{j}^*) + T_{W,i+1}(\hat{j}^*)$

    end for

    $V_{i+1}(j) \leftarrow AVG[TEMP(j, \cdot, \cdot)]$.

    $E_{i+1}(j) \leftarrow AVG[TEMP_{energy}(j, \cdot, \cdot)]$.

    $T_{i+1}(j) \leftarrow AVG[TEMP_{time}(j, \cdot, \cdot)]$

    Where $AVG[TEMP_X(j, \cdot, \cdot)] = \sum_{\ell=0}^{G_R} p_{R,\ell} \sum_{z=0}^{G_{A,\ell}} p_{A,\ell,z} TEMP_X(j, \ell, z)$

  end for

  $\forall \mathbf{s}$, [NM: aer you sure this is the rigth stopping critrerion for VI?? I dont think so because these are cost to go functions and keep growing over time..] calculate the stopping criterion metric $H(\mathbf{s}) = V_{i+1}(\mathbf{s}) - V_i(\mathbf{s})$; $i \leftarrow i+1$.

until $\max_{\mathbf{s}\in\mathcal{S}} H(\mathbf{s}) - \min_{\mathbf{s}\in\mathcal{S}} H(\mathbf{s}) < \delta$. ▷ Value Iteration termination condition

Approximate $\left[g(\nu); \bar{E}; \bar{T}\right] \approx \frac{1}{\pi_{\text{comm}}} \frac{1}{i} [V_i(\mathbf{s}); E_i(\mathbf{s}); T_i(\mathbf{s})]$, [NM: this appears to be a mistake: should be $\left[g(\nu); \bar{E}; \bar{T}\right] \approx \frac{1}{i} [V_i(\mathbf{s})/\pi_{\text{comm}}; E_i(\mathbf{s}); T_i(\mathbf{s})]$,] for some arbitrary $\mathbf{s} \in \mathcal{S}$.

return $O^*(r_U) = O_i(r_U), \forall r_U \in \mathcal{S}_{\text{wait}}$, $U^*(\mathbf{s}) = U_i(\mathbf{s}), \forall \mathbf{s} \in \mathcal{S}_{\text{comm}}$, $g(\nu)$, $\bar{E}$, and $\bar{T}$. ▷ Optimal waiting and communication policies

The one above is the discretized version of the following algo

Algorithm 2 Projected Sub-gradient Ascent: PSGA()

Initialization: k=0; dual variable $\nu_0 \geq 0$; step-size $\{\rho_j = \frac{\rho_0}{(j+1)}, j \geq 0\}$; $g_{-1} = \infty$.
for $k = 0, 1, \ldots$ **do**
  Determine $(O_k^*, U_k^*, g_k, \bar{E}_k, \bar{T}_k) = \mathrm{VITER}\,(\nu_k)$ via Alg. 3.
  if $|g_k - g_{k-1}| < \epsilon_{DI}$; $\bar{E}_k - P_{\mathrm{avg}} \bar{T}_k < \epsilon_{PF}$; $\nu_k |\bar{E}_k - P_{\mathrm{avg}} \bar{T}_k| < \epsilon_{CS}$ **then** ▷ Check KKT optimality conditions
    return: optimal outer policy (O_k^*, U_k^*);
  else
    Update $\nu_{k+1} = \max\left\{\nu_k + \rho_k \left(\bar{E}_k - P_{\mathrm{avg}} \bar{T}_k\right), 0\right\}$; $k \leftarrow k+1$. ▷ Dual variable value update
  end if
end for

Algorithm 3 Value Iteration: $(O^*, U^*, g(\nu), \bar{E}, \bar{T}) = \mathrm{VITER}(\nu)$

Initialization: i=0; value function $V_{W,i}(r_U) = V_{C,i}(r_U) = 0$, total energy cost $E_{W,i}(r_U) = E_{C,i}(r_U) = 0$, and total time cost $T_{W,i}(r_U) = T_{C,i}(r_U) = 0$, $\forall r_U \in [0, a]$, for waiting (W) and communication (C, averaged over GN position and scheduling decision); stop criterion δ.
Inner optimization in waiting states: $\forall r_U \in \mathcal{S}_{\mathrm{wait}}, \forall v_r \in [-V_{\max}, V_{\max}]$, calculate $\ell_\nu^*(r_U; v_r)$ as in (**??**), with minimizer θ_c^*; compute energy cost $e^*(r_U; v_r) = P_{\mathrm{mob}}(\sqrt{v_r^2 + r_U^2(\theta_c^*)^2})\Delta_0$ and time cost $t^*(r_U; v_r) = \Delta_0$.
Inner optimization in communication states: $\forall \mathbf{s} \in \mathcal{S}_{\mathrm{comm}}, \forall \hat{r}_U \in [0, a]$, calculate $\ell_\nu^*(\mathbf{s}; \hat{r}_U, 1) = \min_{\tau \in [0,1]} (1 - \nu P_{avg}) \Delta_\tau(\mathbf{s}; \hat{r}_U) + \nu E_\tau(\mathbf{s}; \hat{r}_U)$, with minimizer τ^* associated to a trajectory $\mathbf{p}_U^*$; compute energy cost $e^*(\mathbf{s}; \hat{r}_U, 1) = E_{\tau^*}(\mathbf{s}; \hat{r}_U)$ and time cost $t^*(\mathbf{s}; \hat{r}_U, 1) = \Delta_{\tau^*}(\mathbf{s}; \hat{r}_U)$.
repeat
  for each $r_U \in [0, a]$ **do** ▷ Outer optimization in waiting states
    $V_{W,i+1}(r_U) \leftarrow \min_{v_r \in [-V_{\max}, V_{\max}]} \left[\ell_\nu^*(r_U; v_r) + e^{-\Lambda' \Delta_0} V_{W,i}(r_U + v_r \Delta_0) + (1 - e^{-\Lambda' \Delta_0}) V_{C,i}(r_U + v_r \Delta_0)\right]$,
    $O_{i+1}(r_U) \leftarrow v_r^*$, where v_r^* is the $\arg\min$.
    $E_{W,i+1}(r_U) \leftarrow e^*(r_U; v_r^*) + e^{-\Lambda' \Delta_0} E_{W,i}(r_U + v_r^* \Delta_0) + (1 - e^{-\Lambda' \Delta_0}) E_{C,i}(r_U + v_r^* \Delta_0)$.
    $T_{W,i+1}(r_U) \leftarrow t^*(r_U; v_r^*) + e^{-\Lambda' \Delta_0} T_{W,i}(r_U + v_r^* \Delta_0) + (1 - e^{-\Lambda' \Delta_0}) T_{C,i}(r_U + v_r^* \Delta_0)$.
  end for
  for each $r_U \in [0, a]$ **do** ▷ Outer optimization in communication states
    for each $r \in [0, a]$, $\psi \in [0, 2\pi)$ $(\mathbf{s} = (r_U, r, \psi))$ **do** ▷ Outer optimization in communication states
      $\hat{V}(\mathbf{s}) \leftarrow \min\{L / R_{GB}(r) + V_{W,i}(r_U), \min_{\hat{r}_U \in [0,a]} \left[\ell_\nu^*(\mathbf{s}; \hat{r}_U, 1) + V_{W,i}(\hat{r}_U)\right]\}$ ▷ Value function given GN position, optimized over scheduling/trajectory
      $U_{i+1}(\mathbf{s}) \leftarrow \hat{r}_U^*$, where $\hat{r}_U^*$ is the $\arg\min$ and ξ^* is the scheduling decision
      $\hat{E}(\mathbf{s}) \leftarrow \xi^* \cdot e^*(\mathbf{s}; \hat{r}_U^*, 1) + E_{W,i}(\hat{r}_U^*)$; $\hat{T}(\mathbf{s}) \leftarrow \xi^* \cdot t^*(\mathbf{s}; \hat{r}_U^*, 1) + T_{W,i}(\hat{r}_U^*)$. ▷ Costs given GN pos., optimized over scheduling/trajectory
    end for
    $V_{C,i+1}(r_U) \leftarrow \int_0^{2\pi} \frac{1}{2\pi} \int_0^a \frac{2r}{a^2} \hat{V}(r_U, r, \psi) \mathrm{d}r \mathrm{d}\psi'$ ▷ Value function in comm states, averaged over GN position
    $E_{C,i+1}(r_U) \leftarrow \int_0^{2\pi} \frac{1}{2\pi} \int_0^a \frac{2r}{a^2} \hat{E}(r_U, r, \psi) \mathrm{d}r \mathrm{d}\psi'$ ▷ Energy cost in comm states, averaged over GN position
    $T_{C,i+1}(r_U) \leftarrow \int_0^{2\pi} \frac{1}{2\pi} \int_0^a \frac{2r}{a^2} \hat{T}(r_U, r, \psi) \mathrm{d}r \mathrm{d}\psi'$ ▷ Time cost in comm states, averaged over GN position
  end for
  $\forall r_U \in [0, a]$, calculate the stopping criterion metric $H_X(r_U) = V_{X,i+1}(r_U) - V_{X,i}(r_U)$, $X \in \{W, C\}$; $i \leftarrow i+1$.
until $\max_{r_U \in [0,a], X \in \{W,C\}} H_X(r_U) - \min_{r_U \in [0,a], X \in \{W,C\}} H_X(r_U) < \delta$. ▷ Value Iteration termination condition
Approximate $g(\nu) \approx \frac{1}{\pi_{\mathrm{comm}}} \frac{V_{W,i}(0)}{i}$, $\bar{E} \approx \frac{E_{W,i}(0)}{i}$, $\bar{T} \approx \frac{T_{W,i}(0)}{i}$.
return $O^*(r_U) = O_i(r_U), \forall r_U \in \mathcal{S}_{\mathrm{wait}}$, $U^*(\mathbf{s}) = U_i(\mathbf{s}), \forall \mathbf{s} \in \mathcal{S}_{\mathrm{comm}}$, $g(\nu)$, $\bar{E}$, and $\bar{T}$. ▷ Optimal waiting and communication policies