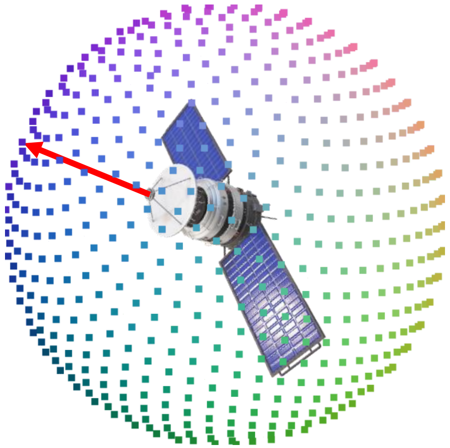
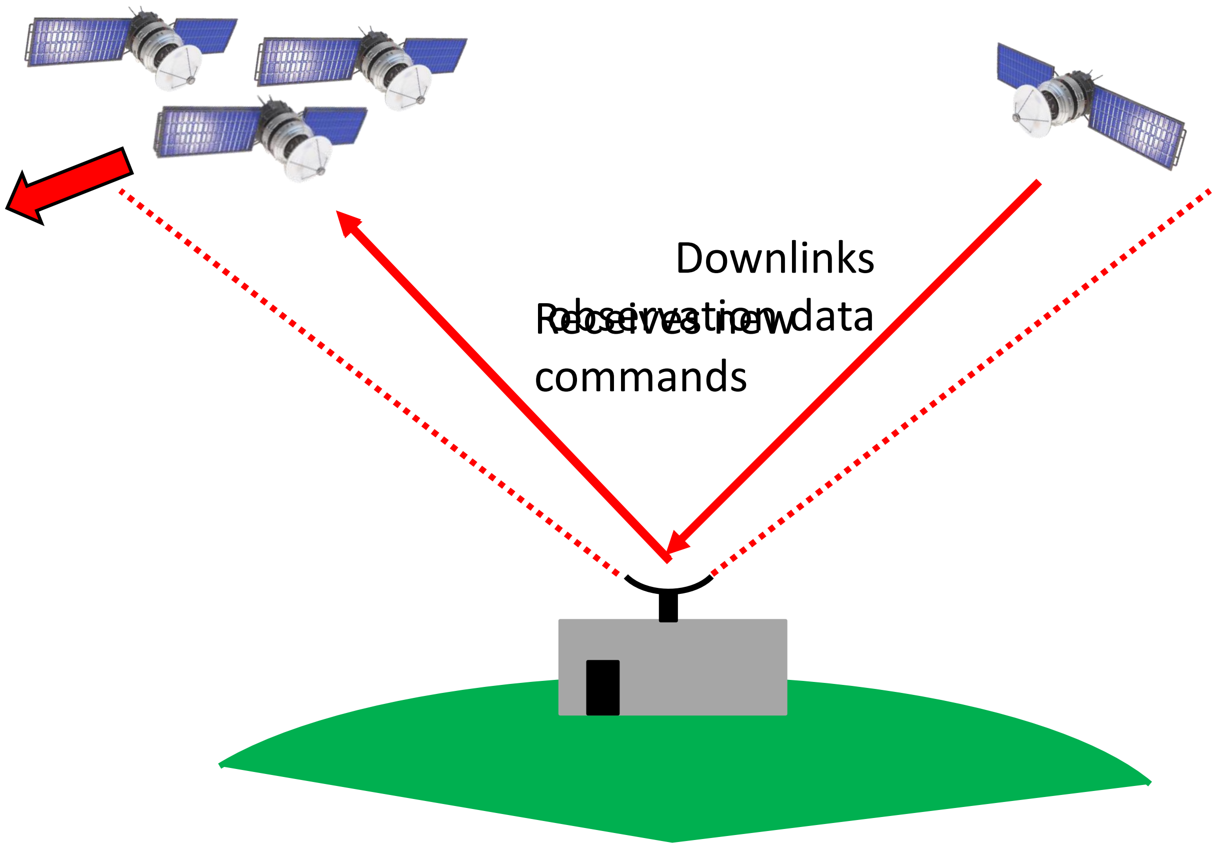
Efficient Strategy Synthesis for Earth-Observing Constellations via MDP Congestion Games



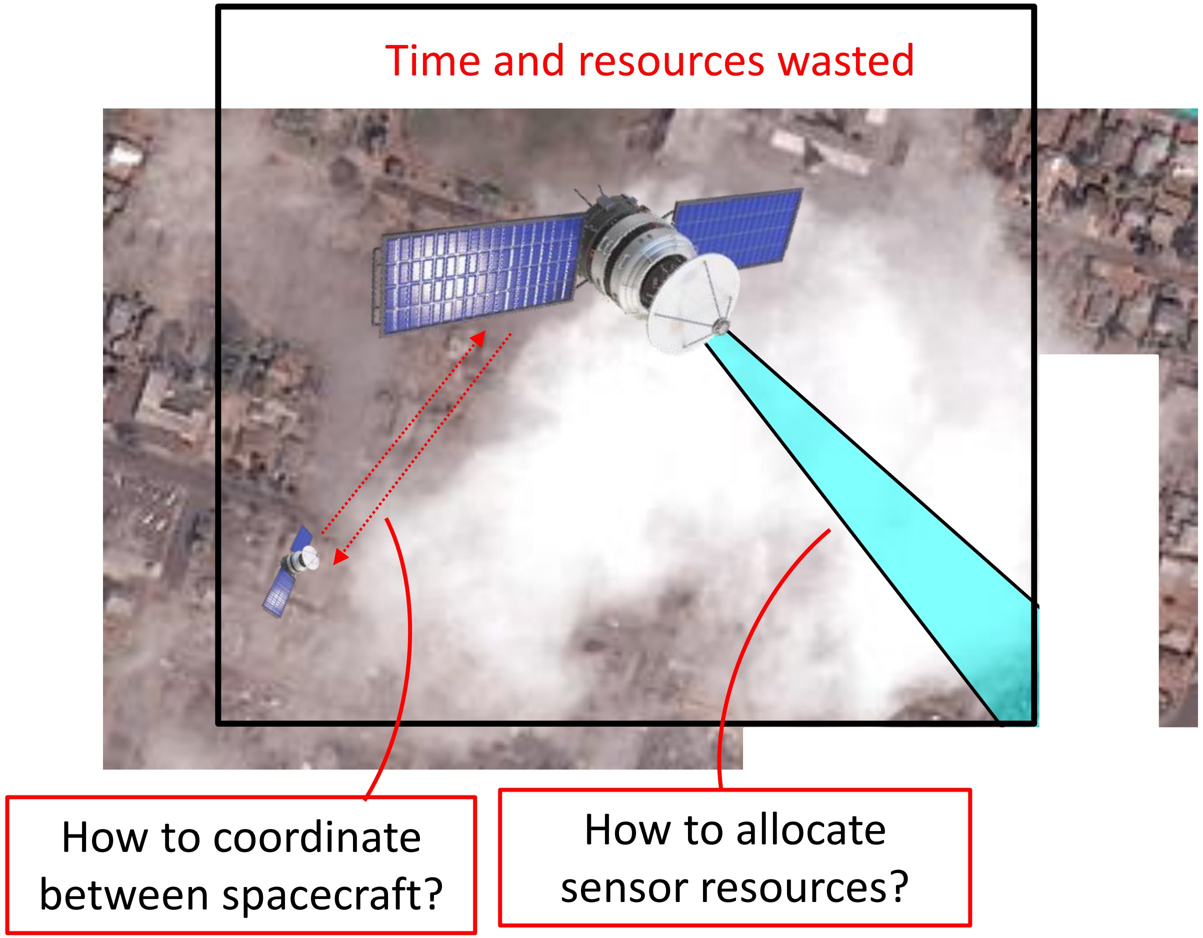
MICHAEL HIBBARD
UNIVERSITY OF TEXAS AT AUSTIN
APRIL 25TH, 2023

aUTonomous
SYSTEMS GROUP

Online Coordination Between Satellites is critical



Redundancy is inefficient!

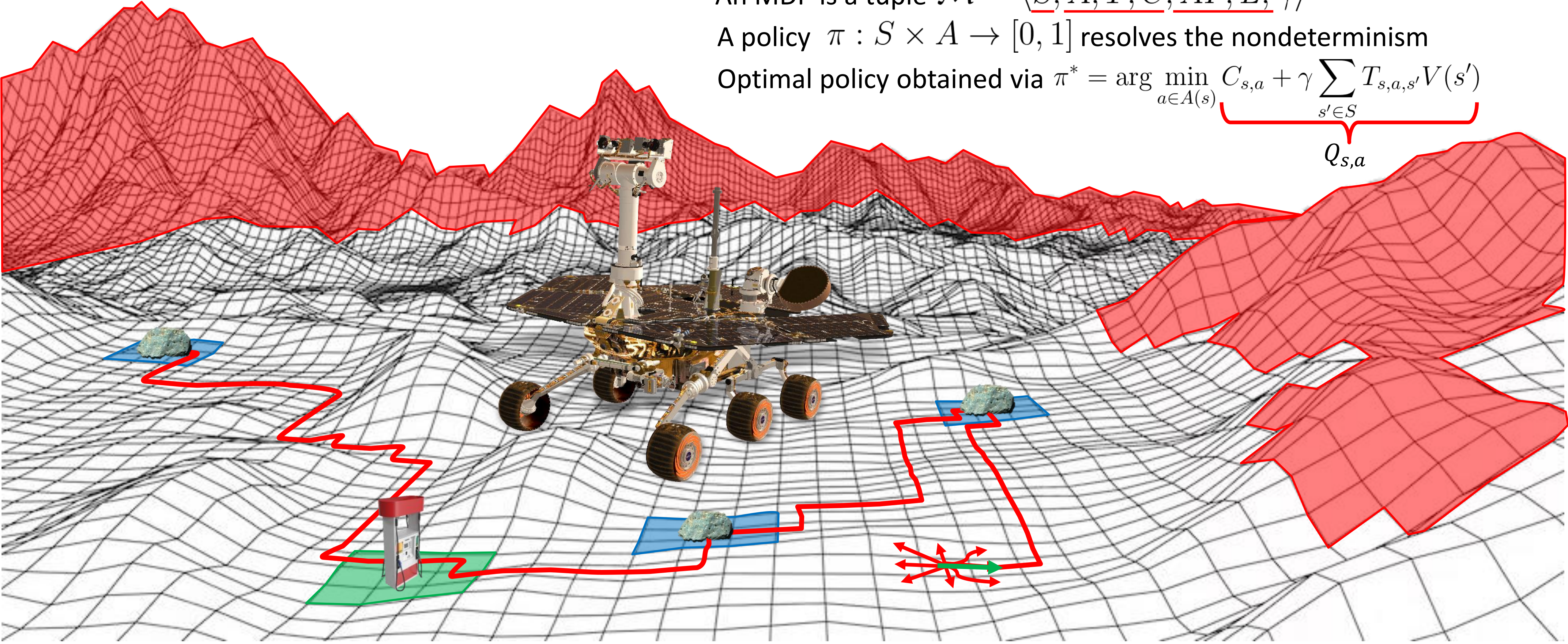


Markov decision processes

An MDP is a tuple $\mathcal{M} = \langle S, A, T, C, AP, L, \gamma \rangle$

A policy $\pi : S \times A \rightarrow [0, 1]$ resolves the nondeterminism

Optimal policy obtained via $\pi^* = \arg \min_{a \in A(s)} \underbrace{C_{s,a} + \gamma \sum_{s' \in S} T_{s,a,s'} V(s')}_{Q_{s,a}}$



Spacecraft attitude representation via quaternions

Euler's Theorem: *Any rotation in 3D space can be expressed as a pure rotation about a single fixed axis*

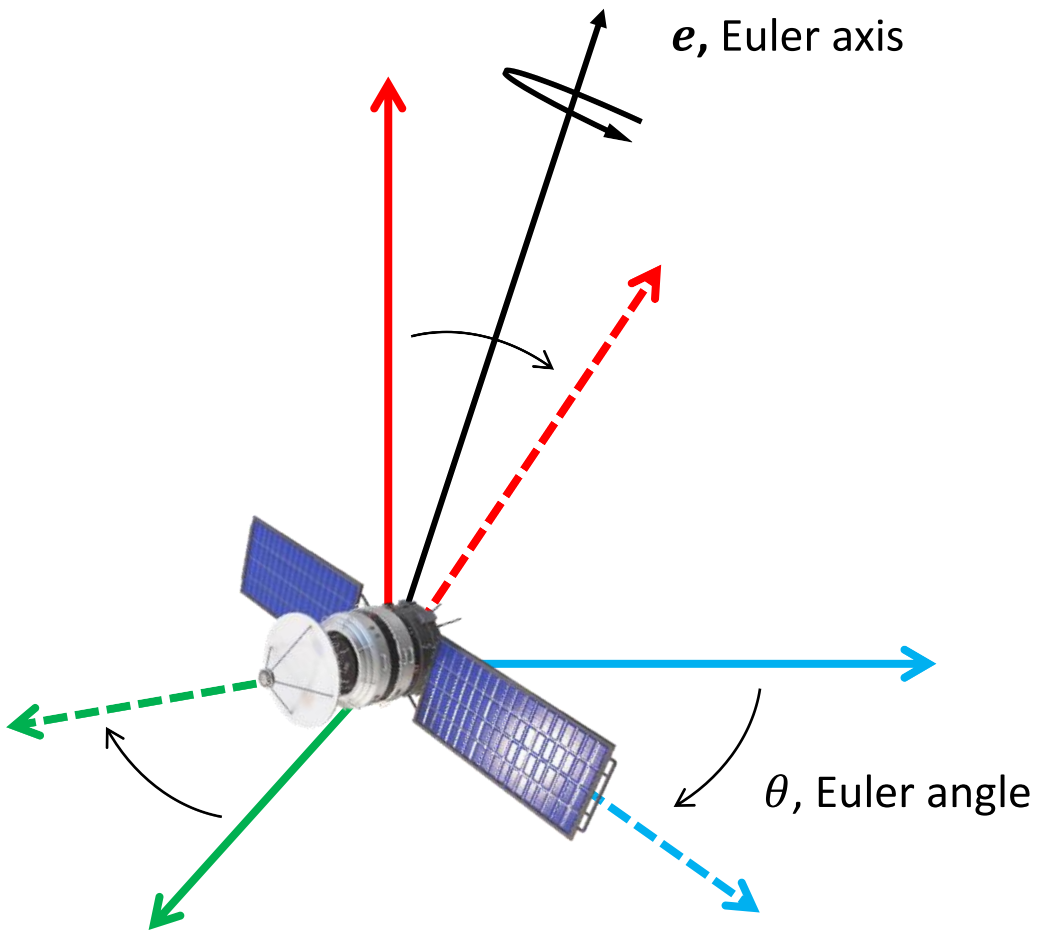
Represent attitude using quaternions:

$$\mathbf{q}(\mathbf{e},\theta) = \begin{bmatrix} \mathbf{e}\sin(\theta/2) \\ \cos(\theta/2) \end{bmatrix}, \quad ||\mathbf{q}||_2 = 1$$

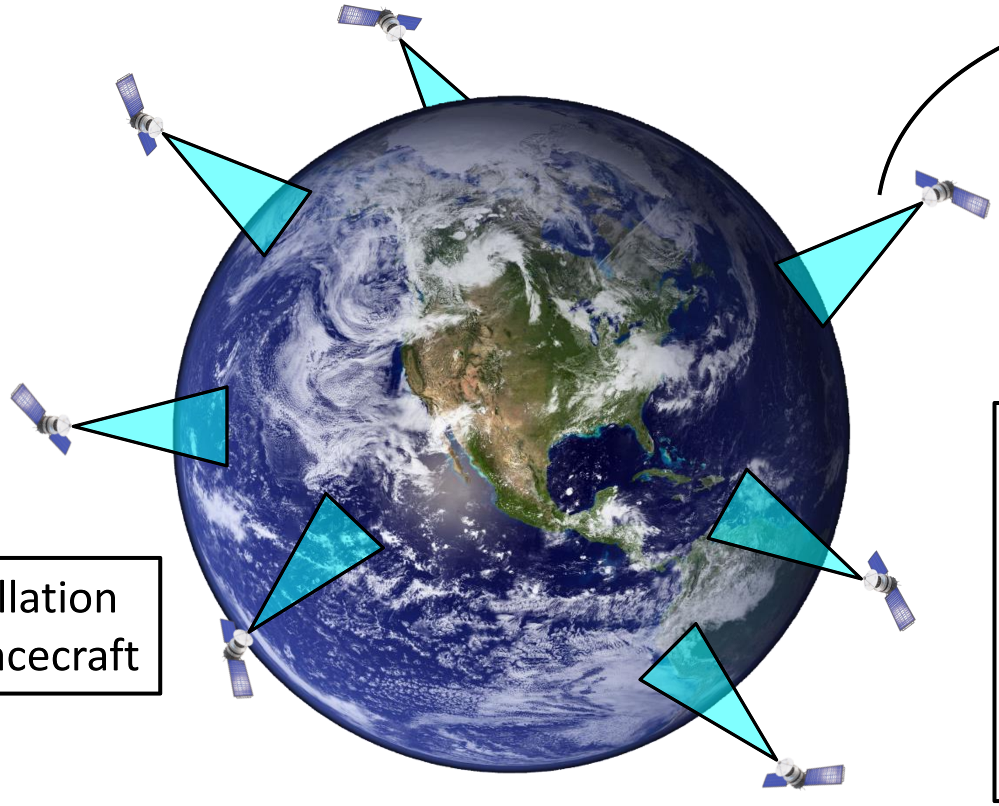
Dynamics and kinematics propagate according to

$$\dot{\mathbf{q}}(t) = \frac{1}{2}\begin{bmatrix} \omega \\ 0 \end{bmatrix} \otimes \mathbf{q}$$

$$J\dot{\omega} = -[\omega\times]J\omega + \mathbf{u} + \mathbf{d}$$



Earth-observing constellations



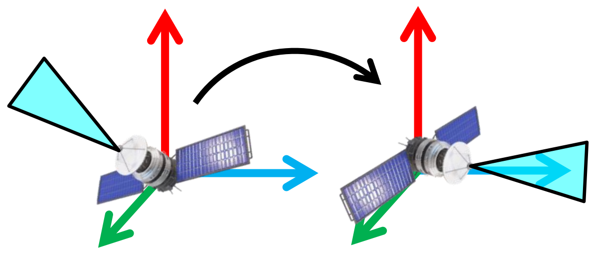
Constellation of N spacecraft

Sensing objective f, seek to solve:

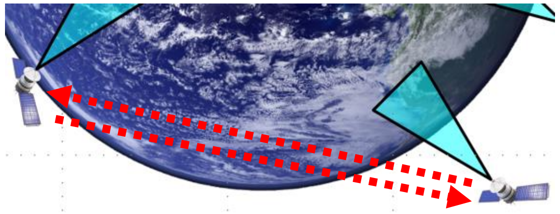
$$\max_{\mathbf{p},\mathbf{q}} \sum_{t=0}^{T} f(\underbrace{p_1(t), q_1(t)}_{\text{Position and attitude}}, \ldots, p_N(t), q_N(t))$$

Subject to constraints on…

Energy usage

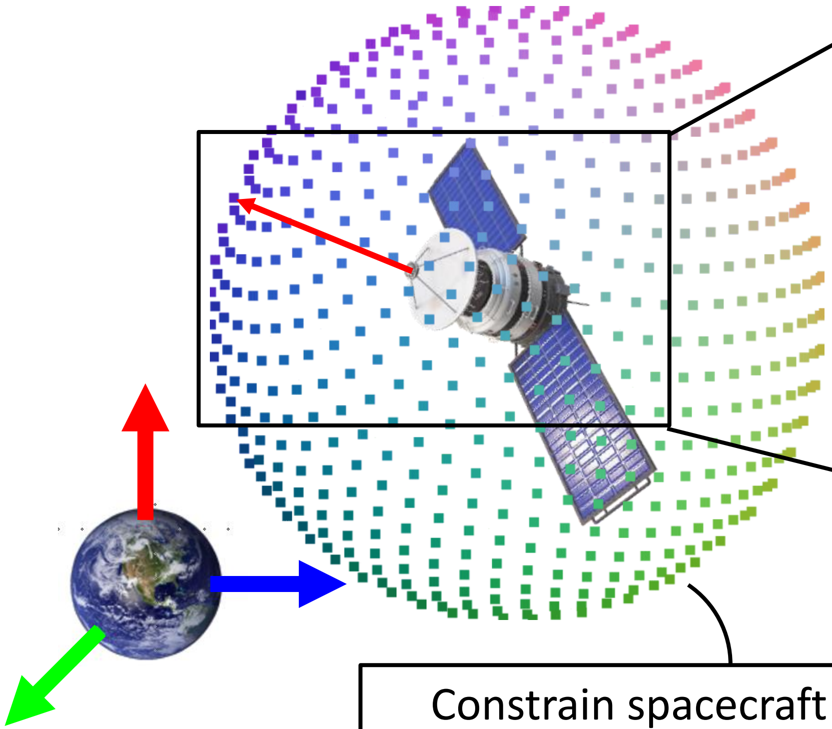


Communication

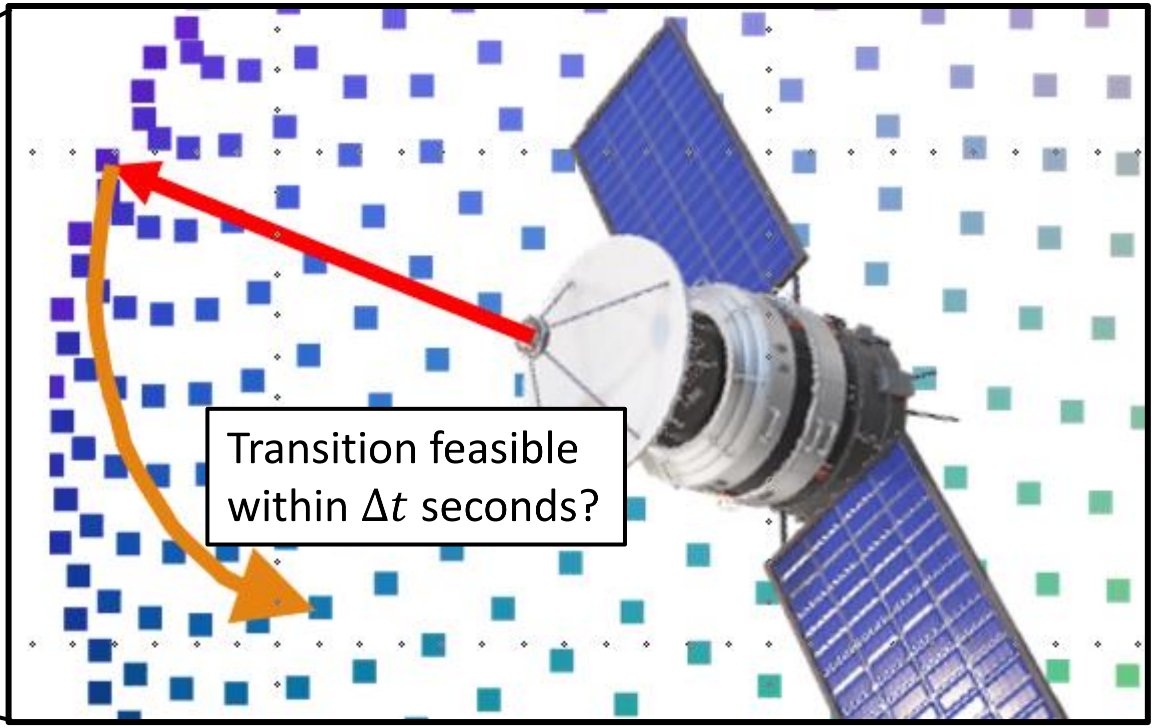


Discretization of the spacecraft environment

Sample a finite set of quaternions corresponding to inertial pointing directions

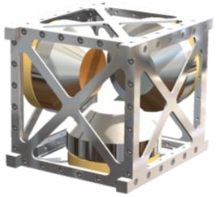


Constrain spacecraft to operate on these directions



Determine feasible transitions and energy costs based on nominal controller

CMG

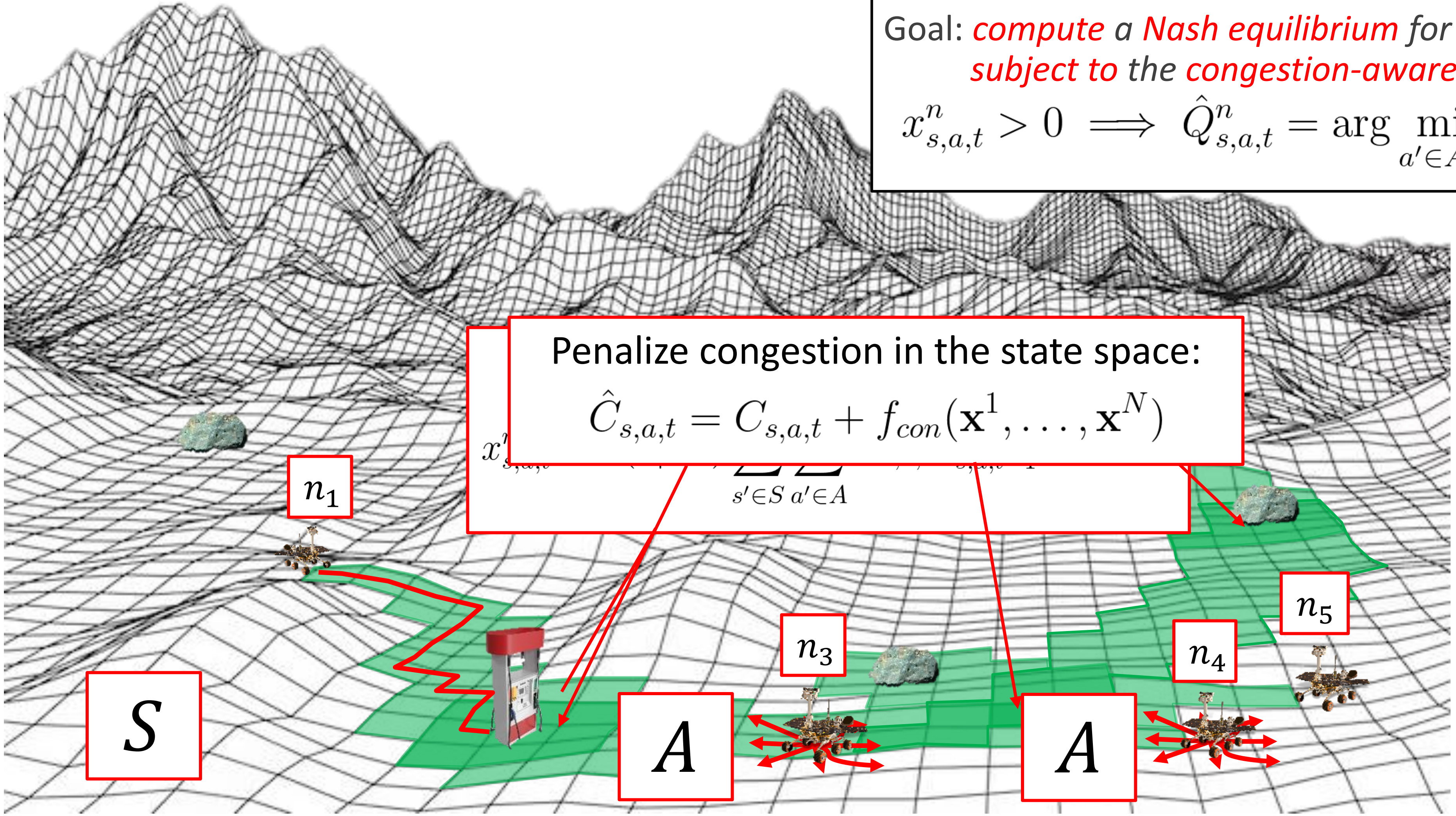


Reaction wheels

MDP congestion games

Goal: *compute a Nash equilibrium for each agent subject to the congestion-aware costs*

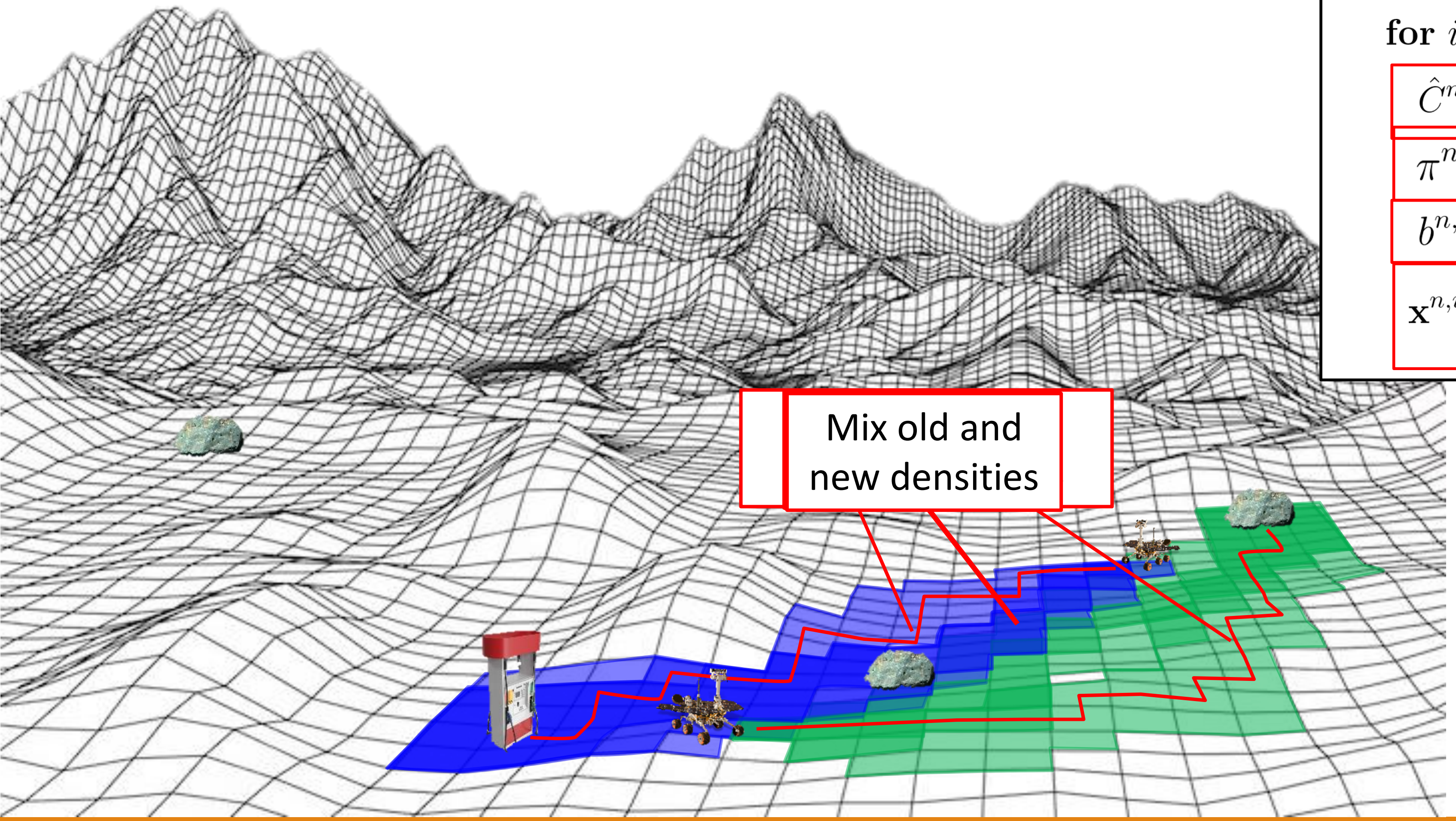
$$x^n_{s,a,t} > 0 \implies \hat{Q}^n_{s,a,t} = \arg \min_{a' \in A(s)} \hat{Q}^n_{s,a',t}$$



MDP congestion game algorithm

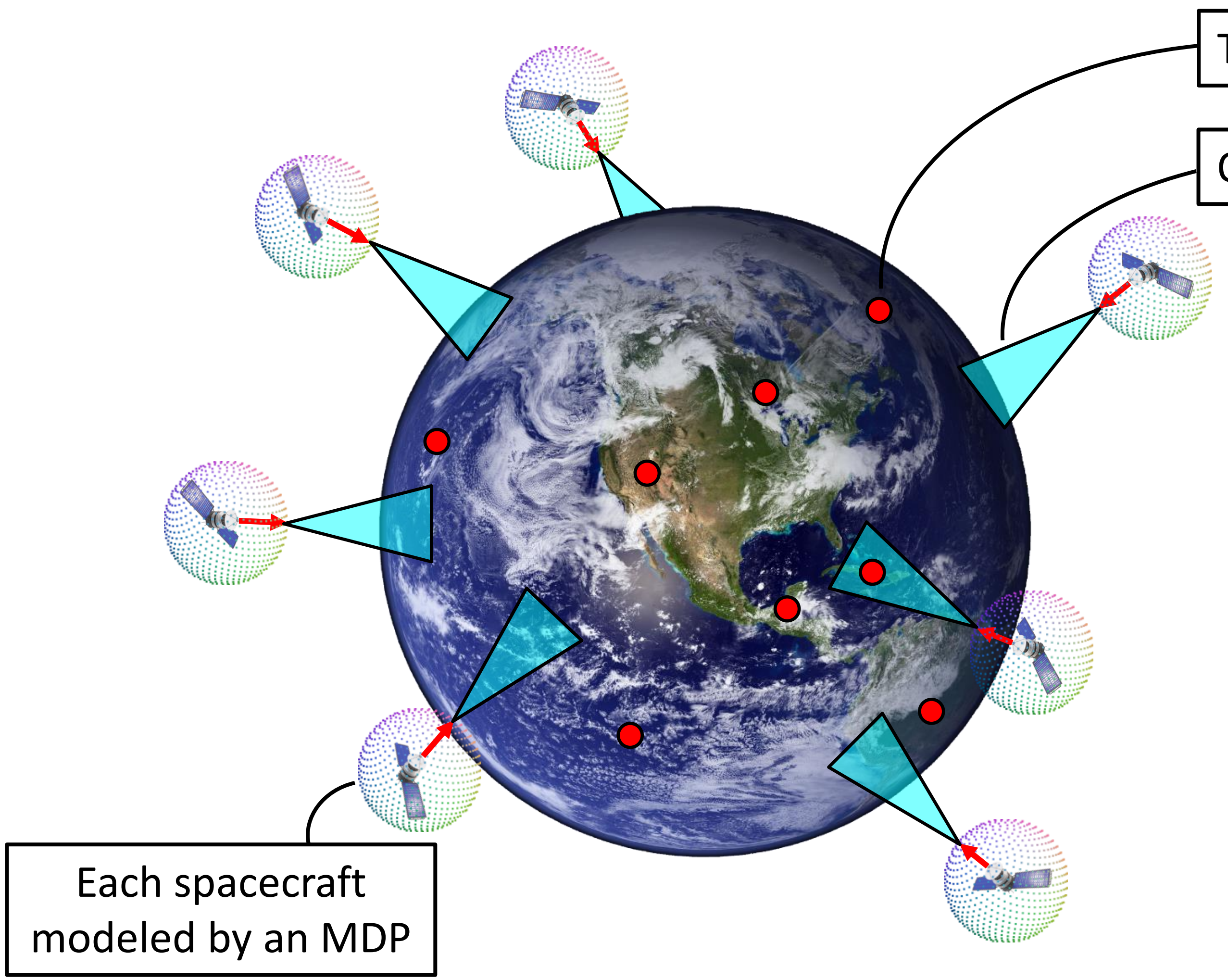
Franke-Wolfe with dynamic programming

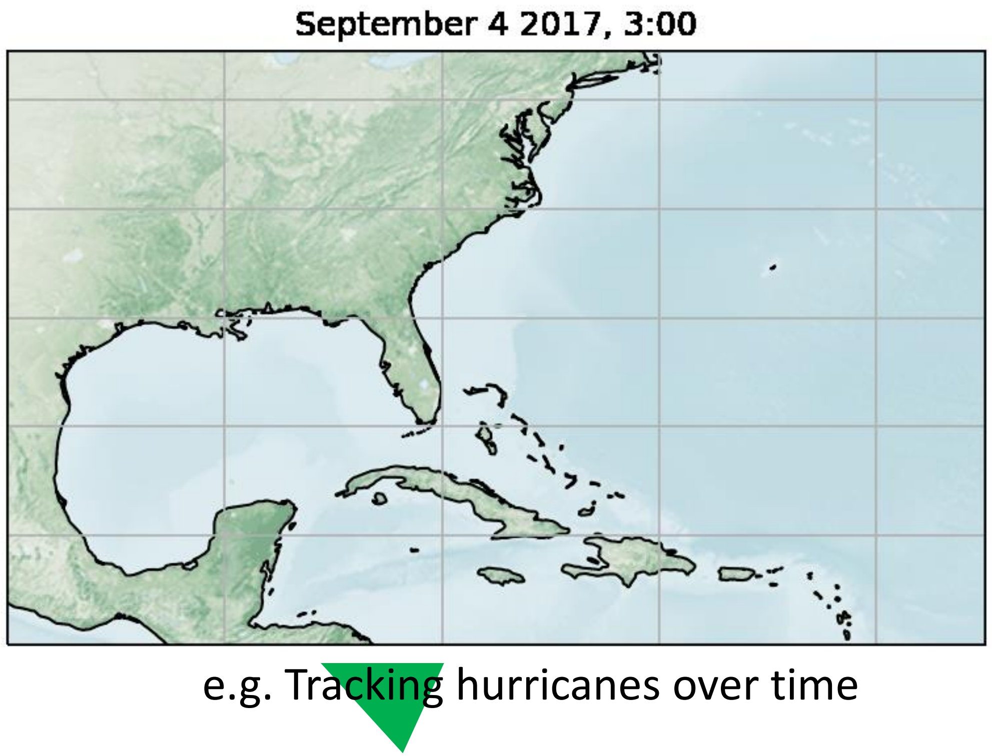
$$
\begin{aligned}
&\textbf{for } n = 1, \ldots, N \textbf{ do}: \\
&\quad \textbf{for } i = 1, 2, \ldots \textbf{ do}: \\
&\quad\quad \hat{C}^{n,i} = f_{con}(\mathbf{x}^{1,i}, \ldots, \mathbf{x}^{N,i}) \\
&\quad\quad \pi^{n} = \text{MDP}(\hat{C}^{n,i}, S, A, T, H) \\
&\quad\quad b^{n,k} = \text{PropDensity}(T, \pi^{n}) \\
&\quad\quad \mathbf{x}^{n,i+1} = \left(1 - \frac{2}{k+1}\right)\mathbf{x}^{n,i} + \left(\frac{2}{i+1}\right) b^{n,i}
\end{aligned}
$$



Under some assumptions on f, obtain globally-optimal Nash equilibrium!

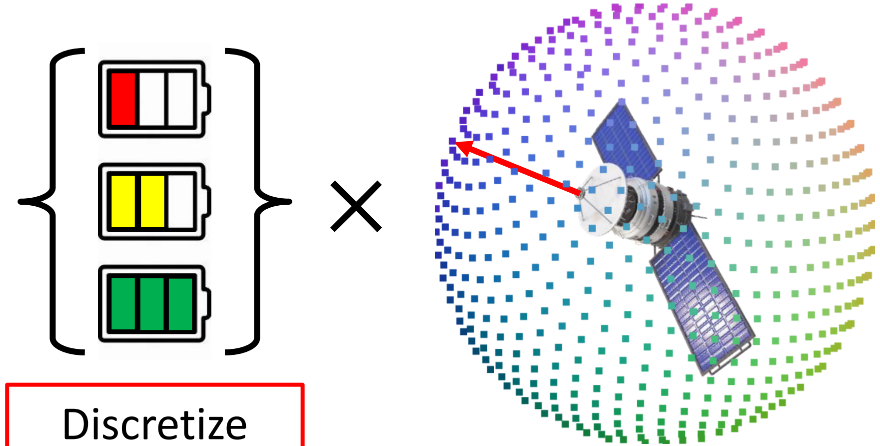
Earth-observing congestion games





e.g. Tracking hurricanes over time

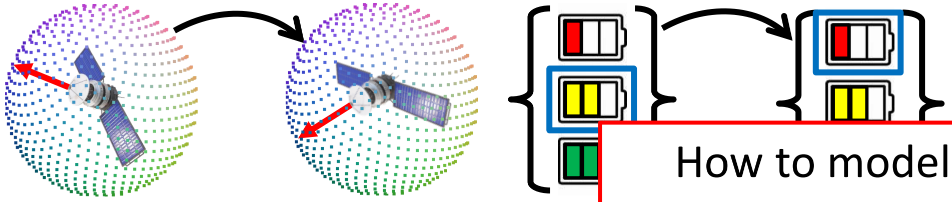
Incorporating energy constraints



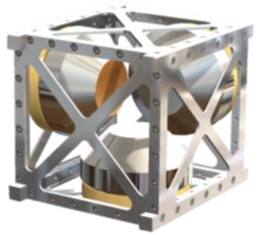
Discretize energy levels

Form product MDP with discretized attitudes

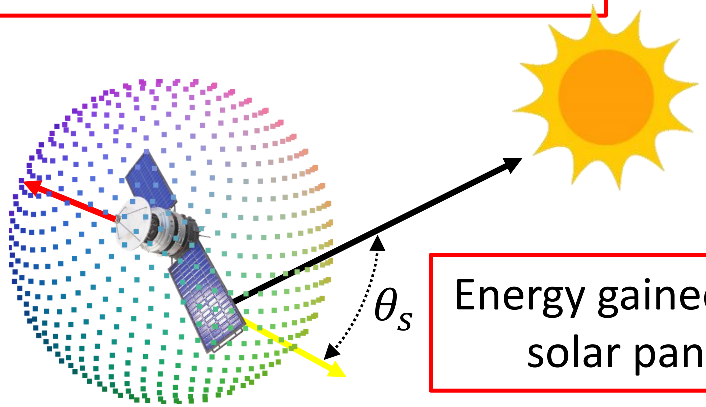
At each time step...



How to model energy transitions?



Energy consumed by controller



Energy gained from solar panels

$$T_{e,e'} = \begin{cases} 1 & e' = \max e'' \quad \text{s.t. } e'' \leq e - E_c + E_s \\ 0 & \text{otherwise} \end{cases}$$

Energy consumed

Energy gained

Incorporating communication constraints

Model communication using directed graph

$$G_{com} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Equals 1 if spacecraft 1 can communicate with spacecraft 7

$$\hat{C}^{n,i} = f(G_{com} \begin{bmatrix} \mathbf{x}^{1,i} \\ \vdots \\ \mathbf{x}^{N,i} \end{bmatrix})$$

for $n = 1, \ldots, N$ **do** :

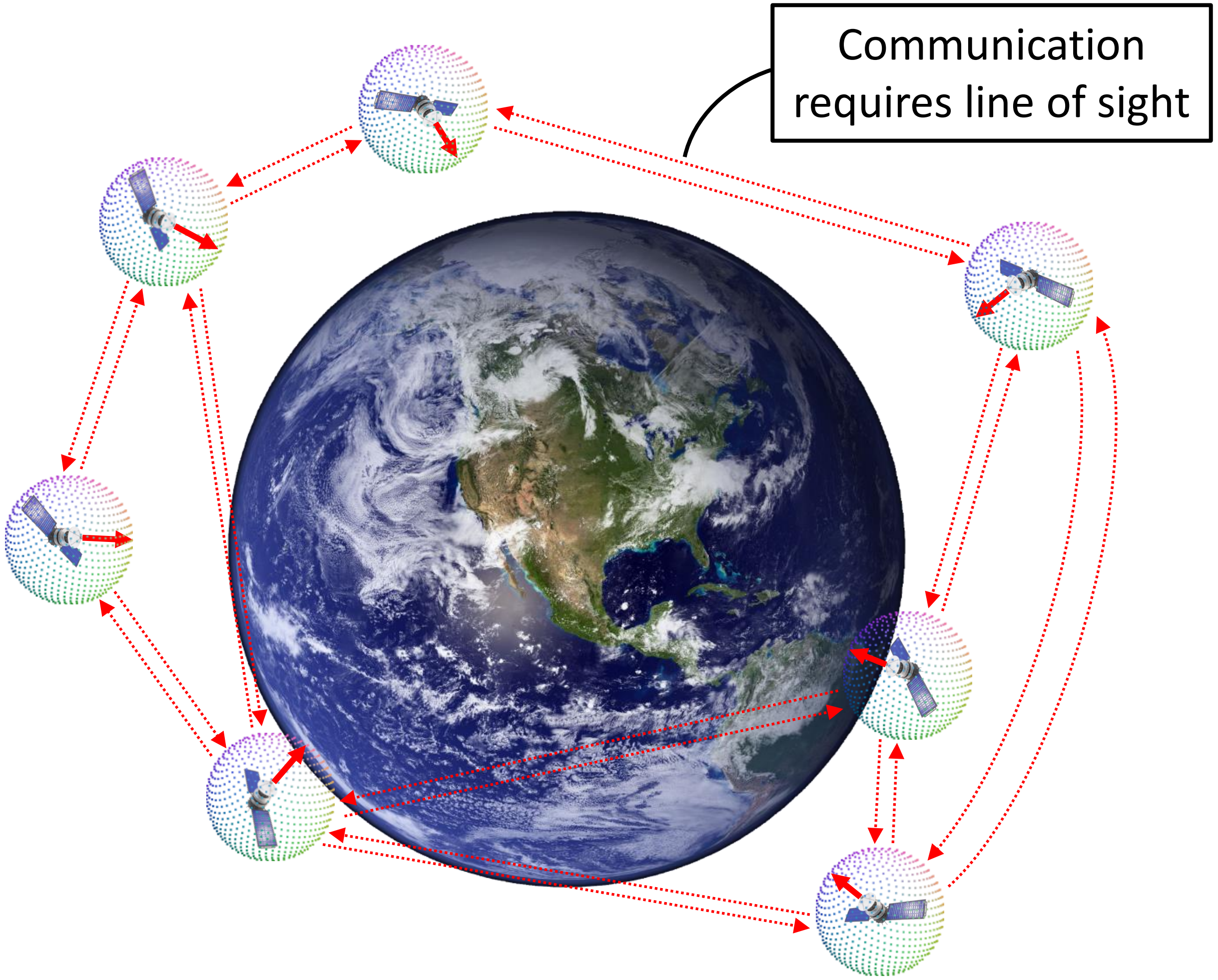
  for $i = 1, 2, \ldots$ **do** :

    $\hat{C}^{n,i} = f_{con}(\mathbf{x}^{1,i}, \ldots, \mathbf{x}^{N,i})$

    $\pi^n = \mathrm{MDP}(\hat{C}^{n,i}, S, A, T, H)$

    $b^{n,k} = \mathrm{PropDensity}(T, \pi^n)$

    $\mathbf{x}^{n,i+1} = \left(1 - \frac{2}{k+1}\right) \mathbf{x}^{n,i} + \left(\frac{2}{i+1}\right) b^{n,i}$



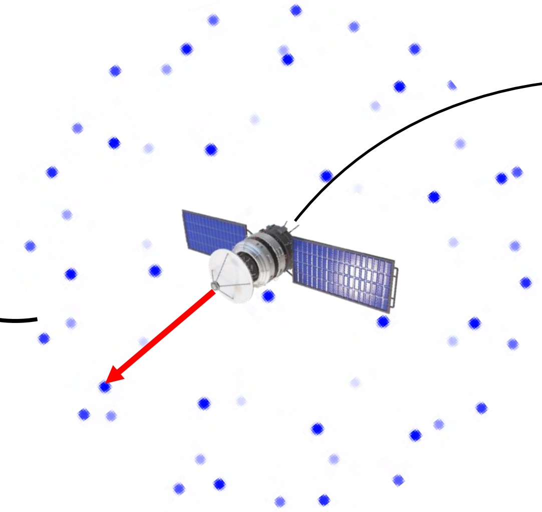
120 spacecraft observing world cities

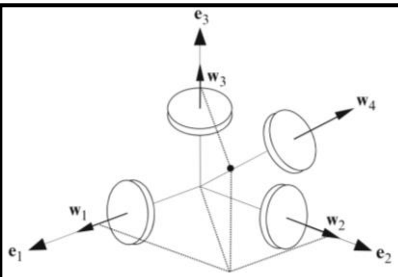
Temperature parameter

Equals one if spacecraft's field-of-view contains city

$$C(p) = -\exp\left(\frac{-\alpha}{H}\sum_{t=1}^{H}\sum_{n=1}^{N}\mathbf{1}^{p}_{x^n_{t,s,a}}\, x^n_{t,s,a}\right)$$

Discretize using 60 inertial pointing directions

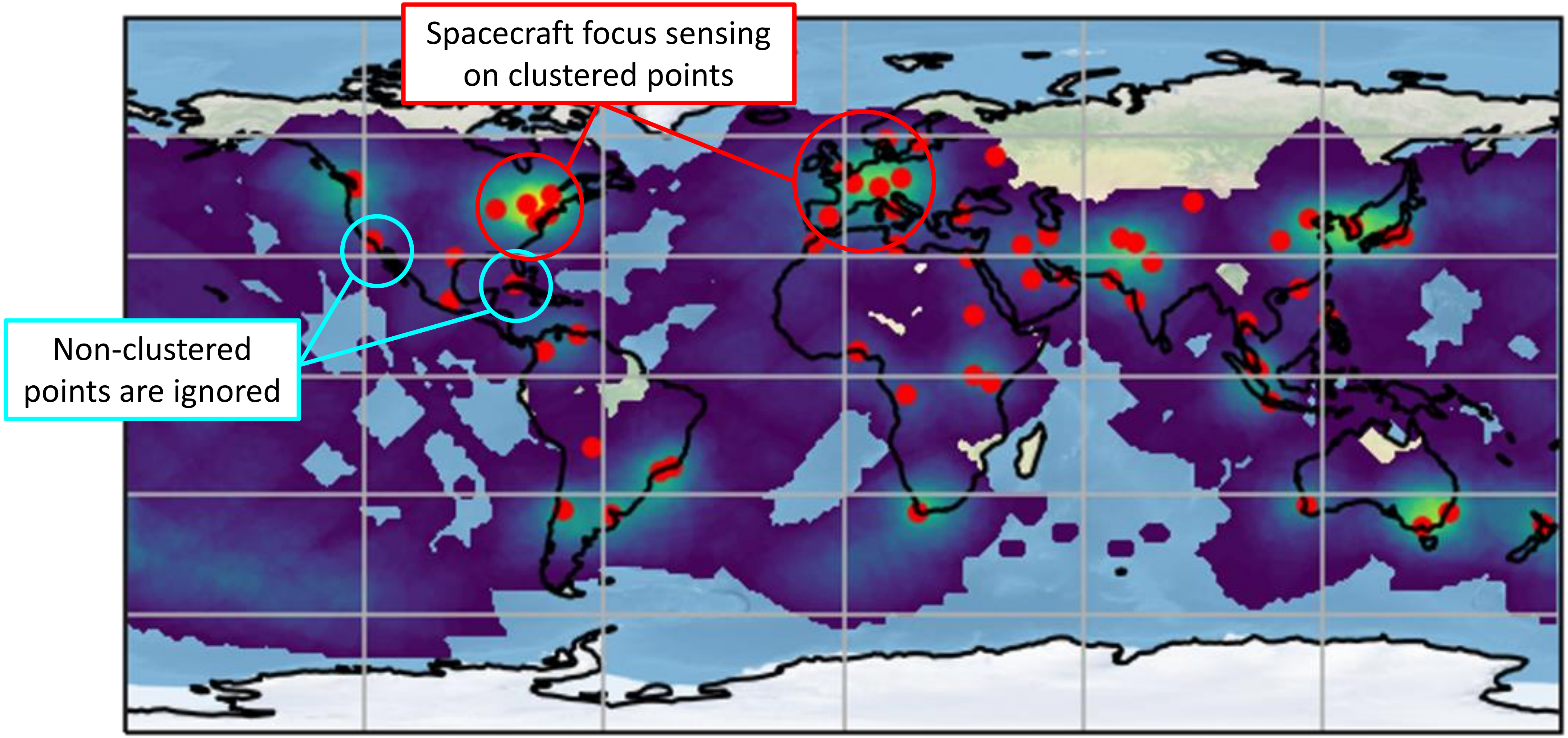




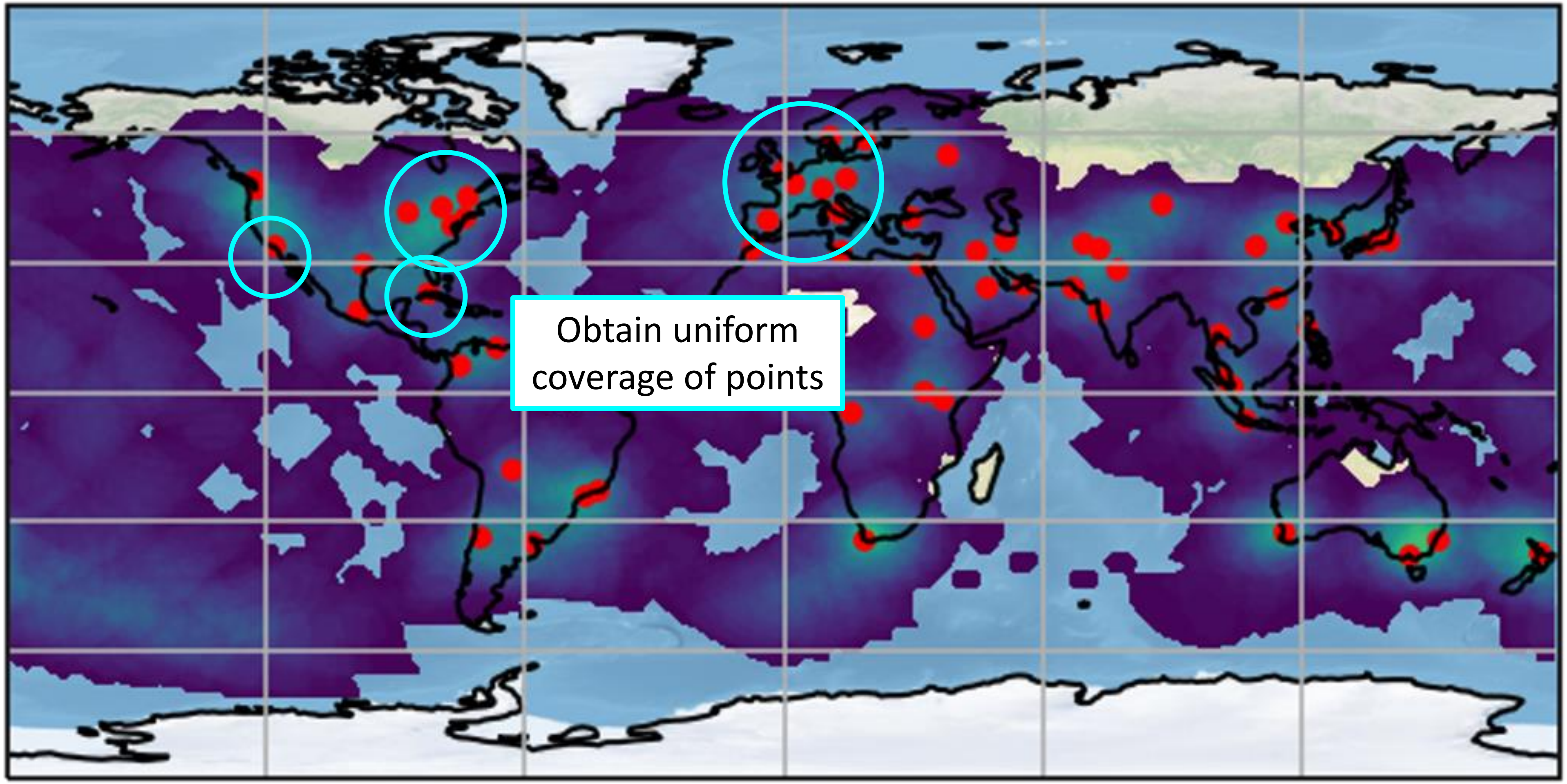
$$W = \begin{bmatrix} 1 & 0 & 0 & 1/\sqrt{3} \\ 0 & 1 & 0 & 1/\sqrt{3} \\ 0 & 0 & 1 & 1/\sqrt{3} \end{bmatrix}$$

Dynamics: reaction wheels with NASA standard configuration, $\Delta t = 60$ seconds

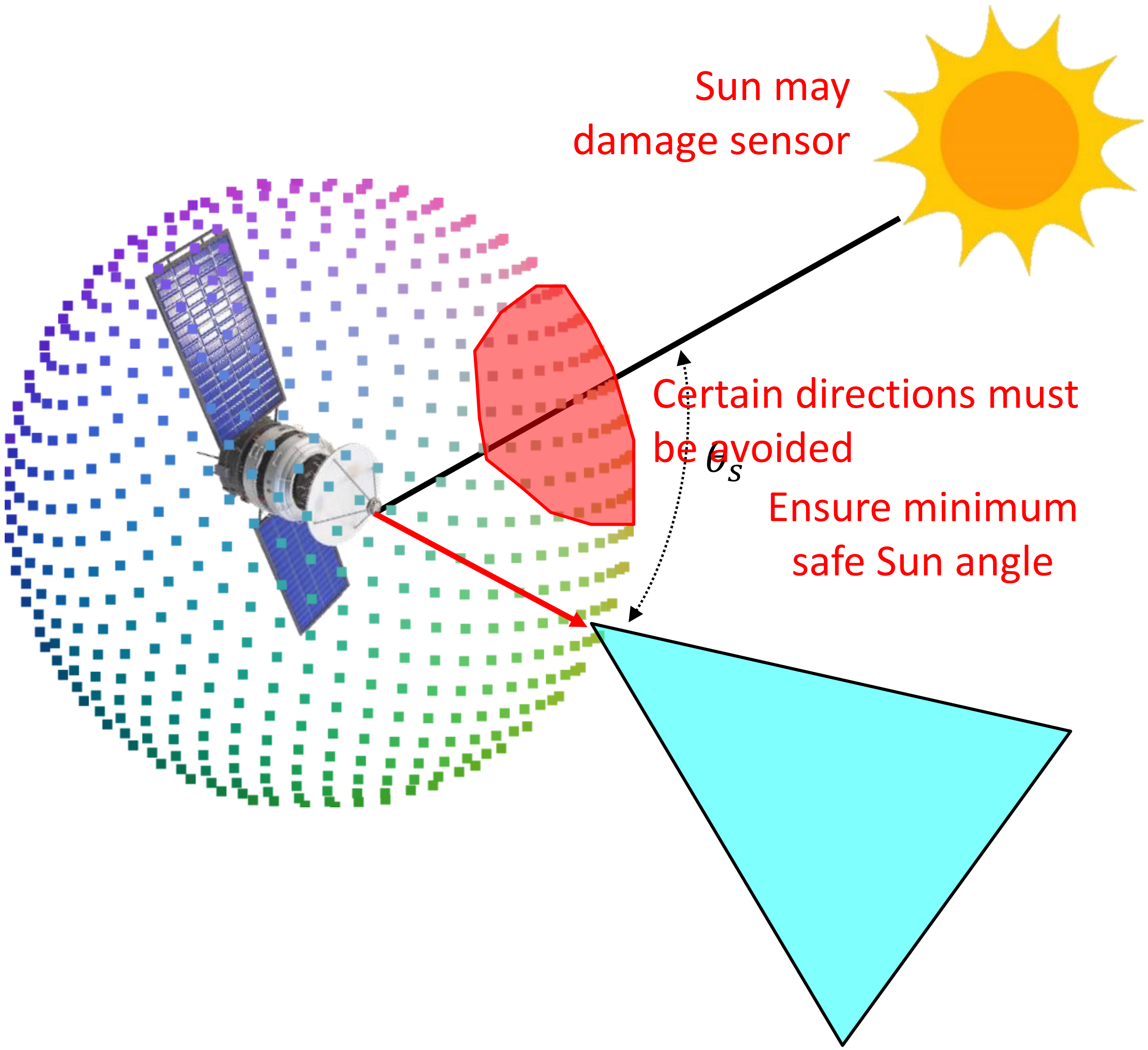
Greedy attitude sequencing ignores points of interest



Congestion costs improve coverage

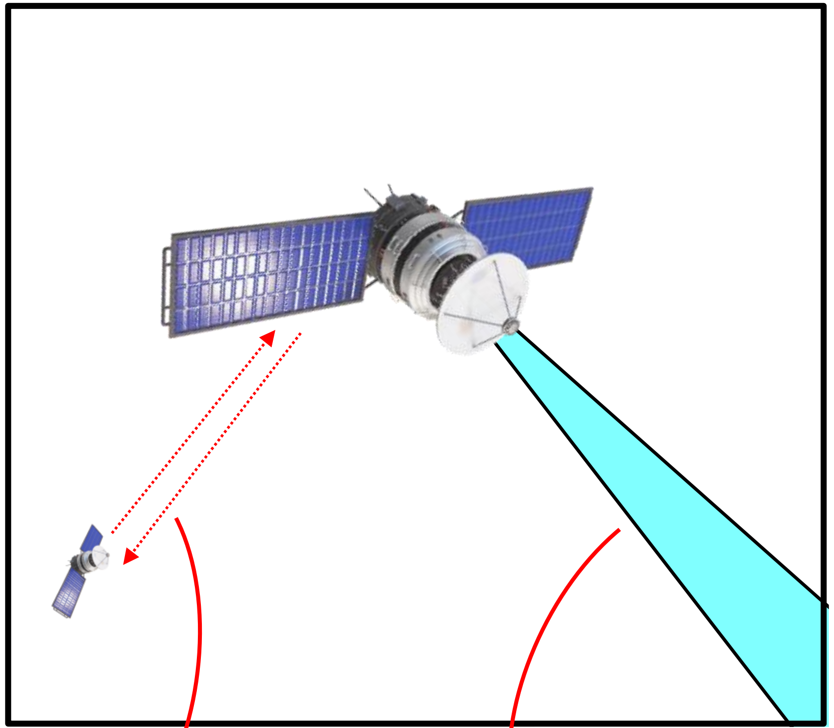


Future work: avoiding keep-out regions



At each time step, only imposes a linear constraint on the residence variables

Congestion games enable efficient strategy synthesis



How to coordinate between spacecraft?

How to allocate sensor resources?