Optimal Multi-Spacecraft Refueling Planning for Cislunar Operations Using Multi-Fidelity Models

Quentin Rommel[1], Michael Hibbard[1], John 'Jack' Chubick[2], Daniel J. Scheeres[3] and Ufuk Topcu[1]

***Abstract*— As space activities expand within the cislunar environment, developing efficient refueling strategies becomes essential for sustaining long-term missions. The multi-spacecraft refueling problem focuses on optimizing fuel consumption to extend mission lifetimes. A multi-fidelity modeling approach addresses varying levels of precision, using the two-body problem for initial trajectory estimation, the circular restricted three-body problem for reference orbits and transfers, and the bicircular restricted four-body problem with solar radiation pressure for realistic station-keeping costs. The refueling problem is formulated as an infinite-horizon Markov decision process (MDP), optimizing the refueling strategy while preventing fuel depletion. The refueling process consists of two phases: first, the refueler transfers into a temporary phasing orbit, followed by a final phasing maneuver to rendezvous with the target spacecraft. We map the available transfers between orbits using a fast sampling technique and select the phasing orbit to minimize fuel consumption within a limited synodic period. The case study samples orbits from the Lyapunov and Halo families around the Lagrange points of the Earth-Moon system, L1, L2, and L3. Simulations demonstrate that the MDP-based planning achieves a 36% reduction in fuel consumption compared to a greedy strategy.**

I. Introduction

Cislunar space, the region between Earth and the Moon, has become a strategic zone for future space exploration, presenting significant challenges for mission management. This strategic zone, often called the "gateway to deep space," is pivotal for future space exploration, serving as a potential staging area for missions to Mars and beyond. Moreover, the ambition to establish a sustained human presence on the Moon, as envisioned by programs like NASA's Artemis, further underscores the importance of cislunar space. A permanent lunar base will require a reliable supply chain for cargo and fuel, making fuel management in this region increasingly complex and vital for long-term mission success [1], [2]. Efficient fuel management will be essential for sustaining spacecraft orbits and executing necessary maneuvers.

In response to these challenges, we optimize the refueling process for multiple spacecraft operating in cislunar space, particularly at key locations such as the Earth-Moon Lagrange points. These points, with their unique dynamical properties, are strategically valuable for various mission objectives, L1 for communication relays and staging, L2 for communications relay for the Moon's far side, L3 for solar observation, and L4 and L5 for potential space stations or interplanetary waypoints [1], [2], [3], [4]. Given the high cost of launching spacecraft from Earth, optimizing resource use is imperative to sustain spacecraft operations without frequent resupply missions [1], [5]. The primary objective is to develop an optimal trajectory planning policy for multi-spacecraft rendezvous operations that minimizes fuel consumption while ensuring that all spacecraft are adequately refueled. By extending the operational lifetimes of these spacecraft, the approach aims to enhance the overall sustainability of space missions in this region [3].

The scenario considered involves a refueling spacecraft stationed at a fuel depot in a stable orbit, tasked with refueling multiple spacecraft located in various orbits within cislunar space. These orbits are sampled from a database [6] of mission-relevant orbits, with their dynamics governed by the circular restricted three-body problem (CR3BP). Each spacecraft in the system consumes fuel at a rate that varies depending on its orbit type and location. The fuel levels of the spacecraft are discretized, and an upper-level controller monitors these levels to ensure that each spacecraft is refueled before its fuel runs critically low. The refueler's mission is to determine the optimal sequence of refueling maneuvers that minimizes the total fuel cost while ensuring that no spacecraft runs out of fuel.

The refueling process involves two distinct phases: transfer and phasing maneuvers. In the transfer phase, the refueler identifies the most fuel-efficient path to reach each target spacecraft's orbit. This phase requires solving a series of trajectory optimization problems aimed at minimizing fuel expenditure. Once the refueler reaches the target orbit, it performs a phasing maneuver to synchronize its trajectory with that of the target spacecraft, aligning their velocities and positions to facilitate the refueling operation. This phase is critical to ensure the refueling operation can be carried out with minimal additional fuel expenditure.

Previous work on cislunar refueling has largely focused on individual spacecraft or short-term missions with immediate, small-scale needs, typically targeting locations such as the Earth–Moon Lagrange points [2]. For example, [5] compares expendable and reusable tanker architectures for sustained lunar and Mars missions, while [7] introduces long-term loitering orbit strategies using CR3BP center manifolds to enable efficient rendezvous, and [8] presents an event-driven network framework that unifies high-thrust and low-thrust

[1]Quentin Rommel, Michael Hibbard, and Ufuk Topcu are with the Department of Aerospace Engineering and Engineering Mechanics, University of Texas at Austin, {quentin.rommel,mwhibbard,utopcu}@utexas.edu.

[2]John 'Jack' Chubick with Westwood High School, jachubick@gmail.com.

[3]Daniel J. Scheeres is with the Ann and H.J. Smead Aerospace Engineering Sciences, University of Colorado Boulder scheeres@colorado.edu.

This work was supported by the Grant AFRL FA9550-23-1-0646.

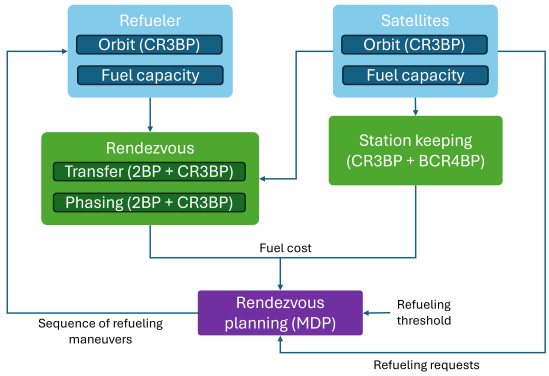


Fig. 1: Refueling process diagram

trajectory models for multi-mission planning. However, these approaches often overlook the challenges of managing multi-spacecraft refueling over extended periods, such as dynamic interactions, non-linear fuel consumption, and cumulative solar perturbations. More recent efforts have started to integrate more realistic dynamics and fuel models into logistics frameworks, yet a comprehensive strategy that fully addresses these complexities remains an active area of research.

We propose an optimized strategy for the multi-spacecraft refueling problem in cislunar space by modeling it as a Markov decision process. This ensures that no spacecraft runs out of fuel while minimizing fuel consumption and determining the most efficient refueling sequence. We integrate models with varying fidelity levels: the two-body problem approximation for the initial guess for trajectory design, the circular restricted three-body problem for efficient orbital dynamics computation, and the bicircular restricted four body problem with solar radiation pressure for accurate station-keeping cost estimation. This approach balances computational efficiency with precision, providing an optimal refueling strategy validated through simulations. Fig. 1 illustrates the relationship between refueling requests, planning, and execution, showing how different models are used for trajectory design, rendezvous planning, and station-keeping to balance simplicity and accuracy.

II. Problem Formulation

We aim to optimize the refueling process for N spacecraft operating in cislunar orbits, with the goal of minimizing fuel consumption while ensuring operational continuity. Each spacecraft follows a distinct orbit, and the refueling spacecraft must service them efficiently. A dedicated refueling spacecraft, positioned in an orbit equipped with a refueling station, is responsible for servicing the system's spacecraft. The problem is complex and can be decomposed into three interrelated subproblems: determining the available transfer trajectories for refueling each spacecraft, selecting phasing maneuvers to rendezvous with the target, and accurately estimating the fuel consumption to stay in orbit for each orbit sampled. The primary goal is to create a robust and efficient refueling strategy that maintains the operational capabilities of all spacecraft. Additionally, insights from the optimization process aim to inform the design of both the spacecraft and refueling infrastructure, aligning them with mission durations, configurations, and operational needs.

A. Infinite-horizon planning using Markov decision processes

The multi-refueling planning problem in cislunar space, characterized by complex mission requirements and resource constraints, is formulated as an infinite-horizon Markov decision process (MDP) [9]. This formulation is particularly well-suited for long-term missions, enabling the optimization of refueling strategies without a fixed endpoint. Given the dynamic and stochastic nature of spacecraft operations in cislunar space, this MDP-based approach provides a robust framework for managing the complexities of multi-refueling planning.

In the context of the refueling optimization problem, the MDP is defined with states (S) representing the position of the refueler in the sampled orbits, its fuel state, and the refueling levels of the spacecraft, as detailed in the next paragraph. Actions (A) correspond to selecting transfer and phasing maneuvers to perform refueling. The initial state ($s_0 \in S$) specifies the starting configuration, including the refueler's position in a distant retrograde orbit (DRO) and the initial fuel levels of all spacecraft. The transition function ($T : S \times A \times S \to [0,1]$) captures the stochastic dynamics of orbital transfers and fuel consumption, governing the probability of transitioning between states after taking specific actions. The reward function ($R : S \times A \to \mathbb{R}$) evaluates the quality of actions based on metrics such as fuel efficiency and operational sustainability, penalizing critical fuel shortages. The discount factor ($\gamma \in [0,1)$) balances the emphasis on immediate versus future rewards, ensuring the development of long-term, sustainable refueling strategies.

The refueling task can be formulated as a set of state-based constraints within the MDP. Tracking the discretized fuel levels quickly becomes intractable as the state space grows exponentially with the number of spacecraft. To limit the size of the problem, we define three key fuel thresholds, l_0, l_1 and l_2, for each spacecraft $i \in \{1,...,N\}$. These thresholds control when refueling must occur and when the refueling process can be considered complete:

- **Critical refueling level:** If the fuel level of spacecraft i drops below the threshold l_0, the corresponding spacecraft requires immediate refueling. Failing to refuel before reaching this critical level leads to a penalty and transition to a failure state.
- **Urgent refueling level:** If the fuel level of spacecraft i is between l_0 and l_1, the refueling is urgent, initiating a "refueling process".
- **Safe refueling level:** If the fuel level of spacecraft i is between l_1 and l_2, refueling is encouraged but not yet critical. It indicates the level at which a spacecraft can be refueled if a "refueling process" is already underway.
- **Operational refueling level:** If the fuel level of spacecraft i is higher than l_2, the spacecraft does not need

refueling.

Each spacecraft fuel is discretized in four levels with values ranging from 0 to 3,

$$s_{fuel,i} = \begin{cases} 0 & \text{if } f_i \leq l_0, \\ 1 & \text{if } l_0 < f_i \leq l_1, \\ 2 & \text{if } l_1 < f_i \leq l_2, \\ 3 & \text{if } f_i > l_2. \end{cases} \tag{1}$$

Once a spacecraft's fuel level exceeds l_2, refueling for that spacecraft is complete. Refueling is successful when all spacecraft have fuel levels above l_2.

The reward function R is designed to favor refueling when a spacecraft's fuel state is at $s_{fuel,i} \leq 1$ and then refuel the other spacecraft at $s_{fuel,i} \leq 2$. The reward for refueling actions is structured as follows,

$$R(s,a) = \begin{cases} M^2 - (\Delta f_r + (1-l_0))^2 & \text{if } s_{fuel,i} = 0, \\ M - (\Delta f_r + (1-l_1))^2 & \text{if } s_{fuel,i} = 1, \\ \alpha M(1-l_1) - (\Delta f_r + (1-l_2))^2 & \text{if } s_{fuel,i} = 2, \\ 0 & \text{if } s_{fuel,i} = 3. \end{cases} \tag{2}$$

Here, $M > 1$ is a tuning parameter, with higher values of M giving more priority to refueling but increasing the risk of depletion. The factor α scales the reward when the fuel level is between l_1 and l_2, allowing better control over intermediate refueling priorities. The term Δf_r represents the fraction of fuel already consumed by the refueler, introducing a penalty to discourage excessive fuel expenditure.

The squared terms in the reward function, $(\Delta f_r + (1-l_j))^2$, serve two purposes. First, it creates a stronger penalty for refueling inefficiencies. This discourages unnecessary refueling trips when the cost outweighs the benefit. Second, squaring the deviation from the threshold levels l_0, l_1, and l_2 ensures a smooth transition in the reward values, preventing abrupt changes that could lead to unstable or erratic refueling decisions. By structuring the penalty quadratically, the function reinforces a more gradual but firm preference for prioritizing spacecraft at the lowest fuel levels while still considering the fuel state of the refueler.

The transition function T incorporates refueling constraints by governing the states' transitions based on fuel levels. The transitions between the orbits are feasible if at least one spacecraft's fuel state is 0 or 1, meaning its fuel level is below l_1. The corresponding transition function is

$$T(s_{orbit,i}, a, s'_{orbit,j}) = \begin{cases} 1, & \text{if } \exists k \in \{1,...,N\} \text{ s.t. } s_{fuel,k} \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

where $i, j, k \in \{1,...,N\}$ represent the indices of different spacecraft.

The objective of the MDP is to find an optimal deterministic policy $\pi : S \to A$ that maps each state $s \in S$ to the action $a \in A$ that should be taken. The optimal policy π is obtained by solving the Bellman equation, which recursively defines the value of each state as the maximum expected reward for taking an action plus the discounted value of the resulting state. We solve it using value iteration, a dynamic programming algorithm that updates the value of each state until convergence, ensuring the computed policy maximizes the cumulative expected reward.

We seek to maximize the expected reward while ensuring all spacecraft are refueled before their fuel levels drop below the critical threshold l_0. The problem can be formulated as the following constrained optimization problem:

$$\max_{\pi} \sum_{s \in S} V(s) \quad \text{s.t.} \quad P(f_i < l_0) \leq \delta \ \forall i \in \{r, 1, ..., N\}. \tag{3}$$

Here, i_r represents the refueler, $i = 1, \ldots, N$ represents the spacecraft that can be refuelled, and δ is the maximum allowed probability that any spacecraft's fuel level drops below the critical threshold. The value function V satisfies the Bellman equation,

$$V(s) = \max_{a \in A} \left[R(s,a) + \gamma \sum_{s' \in S} T(s,a,s') V(s') \right]. \tag{4}$$

with $V(s)$ representing the optimal value function, giving the maximum expected cumulative reward starting from state s and following the optimal policy thereafter.

B. Refueler and refueled orbits

The choice of refueling and target spacecraft orbits plays a key role in optimizing the overall refueling strategy. We select four periodic orbits, referred to as system orbits, from the database [6], focusing on Lyapunov and halo orbits around the Lagrange points L_1, L_2, and L_3 as presented in Fig. 2. These specific orbits were chosen to represent varying orbital environments and stability levels, directly impacting the station-keeping costs The circular restricted three-body problem (CR3BP) defines the environment of the orbits. This model assumes that the mass of the third body (i.e., the spacecraft) is significantly less than those of the other two bodies (i.e., the Earth and the Moon). The first and second body's mass are represented by m_1 and m_2, respectively. The reference frame is centered at the system's barycenter, where the two main bodies are assumed to move in circular orbits around this point. The length, mass, and time units are normalized as per [10], which sets the sum of the primary and secondary masses, the distance between them, and the gravitational constant to 1. Consequently, the mass ratio, μ, is defined as $m_1 = 1 - \mu$ and $m_2 = \mu$. The rotating, non-dimensional equations of motion for the CR3BP are given by

$$\ddot{x} - x - 2\dot{y} = -\frac{(1-\mu)(x+\mu)}{r_1^3} - \mu\frac{(x-1+\mu)}{r_2^3}, \tag{5a}$$

$$\ddot{y} - y + 2\dot{x} = -\frac{(1-\mu)y}{r_1^3} - \mu\frac{y}{r_2^3}, \tag{5b}$$

$$\ddot{z} = -\frac{(1-\mu)z}{r_1^3} - \mu\frac{z}{r_2^3}, \tag{5c}$$

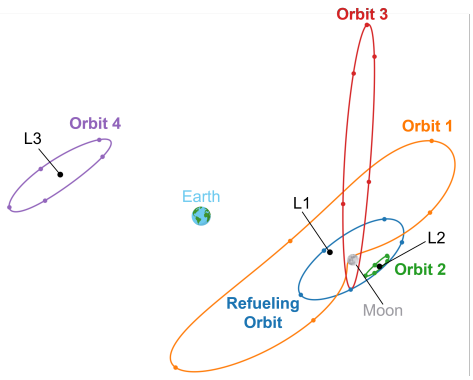


Fig. 2: Selected orbits (system orbits) for the refueling mission in cislunar space. The refueling orbit (blue) is a DRO, while Orbits 1-4 represent various Lyapunov and halo orbits around the L_1, L_2, and L_3 Lagrange points.

TABLE I: Characteristics of the Sampled Orbits

Orbit Name	Orbit Family	Lagrange Point	Jacobi Const. (LU^2/TU^2)	Period (days)	Stability Index
Refueling orbit	DRO	-	2.93	3.28	1.00
Orbit 1	Lyapunov	L1	2.87	7.03	60.1
Orbit 2	Lyapunov	L2	3.16	3.40	657
Orbit 3	Halo (Northern)	L2	3.02	1.97	1.58
Orbit 4	Lyapunov	L3	3.00	6.22	1.67

where

$$r_1 = \sqrt{(x+\mu)^2+y^2+z^2},$$
$$r_2 = \sqrt{(x-1+\mu)^2+y^2+z^2},$$

$p = [x,y,z]$, $v = [\dot{x},\dot{y},\dot{z}]$ and $a = [\ddot{x},\ddot{y},\ddot{z}]$ are respectively the position, the velocity and the acceleration of the spacecraft in the CR3BP reference frame.

The Jacobi constant is the unique solution of the integral of motion. It is defined as

$$C = 2U - \dot{x}^2 + \dot{y}^2 + \dot{z}^2 = 2U - v^2, \tag{6}$$

where v is the total velocity of the spacecraft and U is the effective potential energy, combining gravitational and centrifugal potentials in the rotating reference frame. Lagrange points are positions where gravitational and centrifugal forces balance, allowing a spacecraft to maintain its position. Of the five points (L_1 to L_5), L_4 and L_5 are stable, meaning a spacecraft can remain nearby with minimal adjustments. In contrast, L_1, L_2, and L_3 are unstable, requiring active station-keeping to stay in position.

The refueler spacecraft is positioned in a distant retrograde orbit (DRO), a highly stable orbit encompassing both L_1 and L_2. The DRO was selected for its good stability, which minimizes the need for station-keeping and allows the refueler to efficiently serve spacecraft in both the L_1 and L_2 orbits.

The next step consists of sampling each orbit using the same time step Δt. This process will form the first part of the state space for the MDP. Next, we precisely discretize the fuel level of the refueler and apply the four-level discretization presented in Eq. (1) to the other satellites. We then combine the discretized fuel levels and orbits using a Cartesian product, thereby constructing the complete state space for the MDP.

C. Transfer trajectory design

The transfer computation aims to identify fuel-efficient transfer trajectories between the system orbits, minimizing the overall fuel consumption for refueling maneuvers. First, the different orbits are sampled to define the available transfers with their associated initial and final conditions and costs. These transfers form the basis for the transitions and the rewards in our MDP model.

A double maneuver strategy is adopted in the context of high-level planning for refueling multiple spacecraft. In the first phase, the refueling spacecraft transfers from its current orbit to the phasing orbit. Once the refueling spacecraft reaches the phasing orbit, a phasing maneuver is performed to rendezvous with the targeted spacecraft. The transfers between the orbits are assumed to be instantaneous impulsive maneuvers, with their durations considered negligible compared to the overall transfer time. We compute first the transfers between the system orbits and then add the additional cost and time required for phasing.

The trajectories between sampled points are constructed by solving the Lambert problem, which computes the initial and final velocity vectors required to reach the targeted orbital position within a specified time of flight. In simpler two-body systems, the Lambert problem yields a unique solution. However, in cislunar space, the gravitational influences of the Earth, Moon, and spacecraft introduce complexities that result in multiple feasible transfer trajectories. Since the transfer time is a variable in the Lambert problem, we optimize it to minimize fuel consumption for transfers between X_0 (starting point) and X_f (endpoint). For each transfer time, the fuel cost is evaluated by solving the Lambert problem using a single shooting algorithm, with its outputs, Δv_0 and Δv_f, and then used in the minimization problem

$$\begin{aligned}
\min_{t_f} \quad & J(t_f) = \|\Delta \mathbf{v}_0\| + \|\Delta \mathbf{v}_f\| \\
\text{s.t.} \quad & \dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t)) = \begin{bmatrix} \dot{\mathbf{r}} = \mathbf{v} \\ \dot{\mathbf{v}} = \mathbf{a} \end{bmatrix}, \\
& \mathbf{a} = \begin{bmatrix} x + 2\dot{y} - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3} \\ y - 2\dot{x} - \frac{(1-\mu)y}{r_1^3} - \frac{\mu y}{r_2^3} \\ -\frac{(1-\mu)z}{r_1^3} - \frac{\mu z}{r_2^3} \end{bmatrix}, \\
& \mathbf{X}_0 = \mathbf{X}(t_0) - \Delta \mathbf{v}_0, \quad \mathbf{X}_f = \mathbf{X}(t_f) + \Delta \mathbf{v}_f,
\end{aligned} \tag{7}$$

where t_f is the final transfer time and the optimization variable. The state vector $\mathbf{X}(t)$, comprising the position $\mathbf{r} = [x,y,z]^T$ and velocity $\mathbf{v} = [\dot{x},\dot{y},\dot{z}]^T$, evolves according to the dynamics $\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t))$, with the acceleration $\mathbf{a}$.

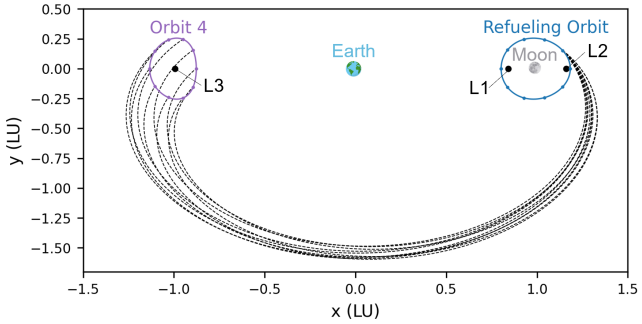


Fig. 3: Transfer from a sampled position of the refueling orbit to orbit 4

The parameters μ, r_1, and r_2 are the same as the ones define in the CR3BP formulation Eq. (5). The initial state $\mathbf{X}_0 = \mathbf{X}(t_0) - \Delta\mathbf{v}_0$ and the final state $\mathbf{X}_f = \mathbf{X}(t_f) + \Delta\mathbf{v}_f$ are adjusted by velocity corrections $\Delta\mathbf{v}_0$ and $\Delta\mathbf{v}_f$, respectively.

We use the single-shooting algorithm because of its efficiency in quickly determining transfer trajectories for all possible combinations of initial and final states from the sampled orbits. This method iteratively refines the trajectory to minimize fuel consumption for transfers. It begins with an initial guess for the trajectory, which is then propagated forward using the system's dynamics. The difference between the propagated final state and the target state, along with the trajectory's final state transition matrix, is used to adjust the initial impulsive maneuver. This process is repeated until the final position error is small enough.

An initial guess is generated by solving the equivalent Lambert problem within the two-body problem to improve convergence in the transfer calculation. The Hill sphere of the relevant astronomical body is calculated, simplifying the CR3BP to a 2BP by identifying the dominant gravitational influence on the spacecraft. In the case of cislunar space, the Hill sphere is determined according to the definition provided in [11], $a_{H-moon} = a\left(\frac{m}{3M}\right)^{\frac{1}{3}} = 61523.869km = 0.157874LU$ where M is the Earth's mass, m is the Moon's mass, and a is the Moon's semi-major axis. If X_0 and X_f are within the Moon's Hill sphere, the Moon is considered the primary body; otherwise, it is the Earth. In this context, Lambert's problem admits a unique solution, with dynamics governed by

$$\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t)) = \begin{bmatrix} \dot{\mathbf{r}} = \mathbf{v} \\ \dot{\mathbf{v}} = -\frac{\mu}{r^3}\mathbf{r} \end{bmatrix}, \tag{8}$$

where $\mathbf{r}$ is the position vector of the spacecraft, $\mathbf{v}$ is its velocity vector, and μ is the standard gravitational parameter of the primary body.

Each orbit is sampled at ten positions, and every possible transfer is computed for each combination of these points. An example of the transfer from a position of the refueling orbit to orbit 4 is presented in Fig. 3.

The transition function incorporates uncertainties in the transfer cost of the MDP. While the transition to the target orbital state is assumed to be deterministic, the fuel transition is treated probabilistically and is defined such that

$$\sum_{j=1}^{m} T(s_{fuel,i}, a_{trans}, s'_{fuel,j}) R(s_{fuel,i}, a_{trans}, s'_{fuel,j}) = \alpha C_{trans}.$$

The reward function is defined as $R(s_{\text{fuel},i}, a_{\text{trans}}, s'_{\text{fuel},j}) = 1 - (s_{\text{fuel},i} - s'_{\text{fuel},j})/s_{fuel,max} = C_{\text{trans}}$, where $s_{\text{fuel},i}$ and $s'_{\text{fuel},j}$ represent the fuel levels before and after the transfer, with $i, j \in \{1, \ldots, m\}$, corresponding to the m discretized fuel levels of the refueler. $\alpha > 0$ is a coefficient to account for uncertainties from the simplified model and stochastic effects in the space environment.

D. Phasing maneuvers for efficient rendezvous

Once the phasing orbit is reached, a phasing maneuver is performed. A phasing maneuver adjusts the refueler's orbit to match the position and velocity of the target spacecraft.

The fuel cost is estimated using the Jacobi constant of the phasing orbit and the spacecraft's refueled orbit in order to select an optimal orbit for phasing. We use Eq. (6) and assume that, since the two orbits belong to the same family and have similar parameters, the potential $U_p \approx U_t$, where U_p is the potential of the phasing orbit and U_t is the potential of the target orbit. The cost of phasing can be estimated using $C_p - C_t = 2U_p - v_p^2 - 2U_t + v_t^2 \approx v_t^2 - v_p^2 \Rightarrow \Delta v = (v_t - v_p) \approx \frac{C_p - C_t}{v_t + v_p}$. The worst case is assumed, so v_p and v_t are considered the minimum velocities of each orbit.

The second criterion for orbit selection is the time required for phasing, which is calculated using the synodic period, $T_{\text{syn}} = \left|\frac{1}{T_t} - \frac{1}{T_p}\right|$. The synodic period represents the time it takes for two spacecraft to "lap" each other in their respective orbits. The phasing time is the duration needed to eliminate the initial phase difference, $\Delta\theta_0$, such that both spacecraft align in their orbits in $t_{\text{phasing}} = \frac{\Delta\theta_0}{\frac{2\pi}{T_{\text{syn}}}}$.

The phasing orbits are presented in Table II. No orbit was selected for the Lyapunov around L3, as all the orbits from this family share remarkably similar periods. Therefore, the synodic period tends to be in the order of years, making it completely impractical. Phasing orbits closer to the target orbit have typically lower fuel costs due to their similar Jacobi constants. However, these orbits may result in a longer synodic period, increasing the time required for phasing maneuvers. Conversely, selecting a more distant orbit may reduce the time needed for phasing but increase fuel costs. Careful selection of phasing orbits ensures that the refueler and target spacecraft can rendezvous within a manageable synodic period and with minimal fuel consumption. Optimizing these maneuvers is crucial in reducing the overall fuel expenditure and maximizing mission efficiency.

After selecting the phasing orbit, we calculate the phasing maneuvers using the same formulation from Problem (7), sampling the system orbits and their corresponding phasing orbits. In this context, the optimization variables are $\mathbf{X}_0$, representing the initial position of the refueler on the phasing orbit, and $\mathbf{X}_f = \mathbf{X}_{i,j}(t_0 + kt_{\text{phasing}})$, representing the final position shared by both the targeted spacecraft and the refueler. Here, i denotes the index of the spaceship to refuel,

TABLE II: Phasing orbit characteristics

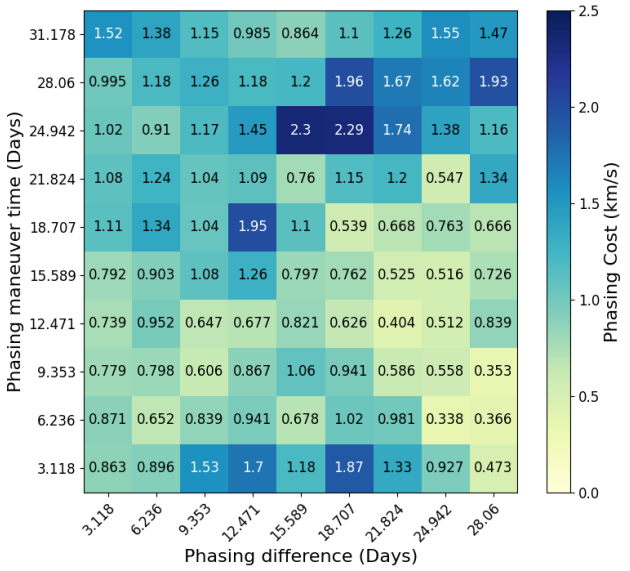
Phasing orbit	Lagrange point	Jacobi const. (LU^2/TU^2)	Δv (km/s)	Phasing orbit period (days)	Synodic period (days)
DRO	-	3.08	0.156	5.13	62.7
Lyapunov	L1	2.96	0.160	23.3	92.2
Lyapunov	L2	3.04	0.470	17.8	97.2
Halo (Northern)	L2	3.03	0.156	7.98	12.4
Lyapunov	L3	-	-	-	-

and j corresponds to its discretized orbital state. The phasing time, t_{phasing}, can either be a constraint if the refueler shares the target's orbit for phasing or an additional optimization variable in cases where a different phasing orbit is selected. A Δv cost heatmap comparison of those two cases for orbit 1 is shown in Fig. 4. Using a phasing orbit enables the selection of the global minimum cost, as the phasing between the refueler and the target can be actively controlled. If the same orbit as the target is used, the cost is estimated by the highest minimum value across the different columns. By selecting a different phasing orbit, the cost can be reduced significantly. In the example shown in Fig. 4, the cost is reduced by a factor of 3, decreasing from 739 m/s (first column, seventh row of Fig. 4a) to 263 m/s (fifth column, ninth row of Fig. 4b). This approach is applied to all orbits except the one around L3, which lacks alternative orbits within the same family with varying periods.

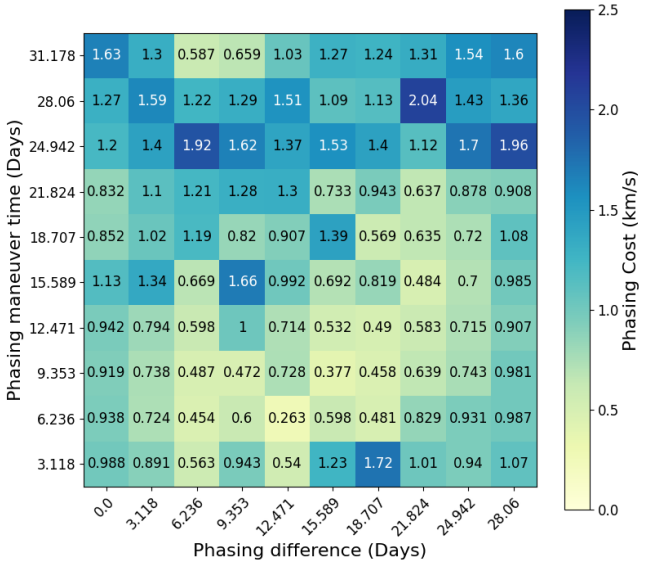
E. Station-keeping cost estimation

Station-keeping is vital in maintaining spacecraft in stable orbits over long-term missions, particularly in cislunar space where gravitational perturbations from the Earth, Moon, and Sun must be managed. Efficient station-keeping ensures that spacecraft remain in position with minimal fuel consumption. We estimate the station-keeping fuel cost using a discrete control approach, assuming impulsive maneuvers similar to the transfer computations. To account for the Sun's gravitational influence, the bicircular restricted four-body problem (BCR4BP) with the Sun-out-of-the-plane from [12] is used. In this model, the dynamics are governed by the Earth, Moon, and Sun, with the Sun moving out of the orbital plane. This problem assumes that the primary bodies, Earth, Moon, and Sun, are in circular orbits around their respective barycenters, with the fourth body (the spacecraft) being influenced by their gravitational fields but not exerting a significant influence on them. The BCR4BP extends the traditional CR3BP by adding the perturbative effect of the Sun, which is particularly relevant for cislunar space, where long orbital periods make solar perturbations more significant.

Additionally, we incorporate solar radiation pressure (SRP), which models the effect of solar photons impacting the spacecraft. This force can cause gradual drifts in the spacecraft's orbit over time, especially for spacecraft in cislunar space, where small forces can have significant



(a) Heatmap phasing cost using the target orbit (orbit 1)



(b) Heatmap phasing cost using a phasing orbit (orbit 1)

Fig. 4: Heatmap phasing cost (orbit 1). (a) Target orbit for phasing, (b) phasing orbit for phasing.

cumulative effects. The SRP model follows the formulation in [10], where the force is proportional to the surface area of the spacecraft exposed to the Sun, the reflectivity of the spacecraft's material, and the inverse square of the distance from the Sun.

The station-keeping method applied here is a variation of the multiple impulses optimization technique described in [13]. This technique divides the orbit period into k segments, with an impulsive maneuver executed at each node i for $i = 1$ to k. Unlike the standard formulation, the initial position is not treated as a decision variable in this approach. Instead, the cost function accounts for the difference between the final and the corrected initial velocities. The objective is to

TABLE III: Cost Analysis for station-keeping of the sampled orbits

Orbit family	Lagrange point	No. of impulses	Cost per orbit (m/s)	Orbits per year	Cost per year (m/s)
Lyapunov	L1	13	7.8	12	93.6
Lyapunov	L2	12	5.2	25	130
Halo (Northern)	L2	3	2.1	42	88.2
Lyapunov	L3	3	4.3	14	60.2

minimize total fuel consumption for station-keeping while ensuring the spacecraft returns within a certain error margin to its initial position and velocity, resulting in the following optimization problem,

$$
\begin{aligned}
&\min_{\mathbf{X}_{opt}} \quad J(\mathbf{X}_{opt}) = \sum_{i=0}^{k-1} |\Delta \mathbf{v}_i| + \frac{1}{2}\mathbf{e}_f^T H \mathbf{e}_f \\
&\text{s.t.} \quad \dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t)) = \begin{bmatrix} \dot{\mathbf{r}} = \mathbf{v} \\ \dot{\mathbf{v}} = \mathbf{a}_{grav} + \mathbf{a}_{srp} \end{bmatrix}, \\
&\mathbf{a}_{grav} = \begin{bmatrix} x + 2\dot{y} - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3} - \frac{m_s(x-x_s)}{r_{s3}^3} \\ y - 2\dot{x} - \frac{(1-\mu)y}{r_1^3} - \frac{\mu y}{r_2^3} - \frac{m_s(y-y_s)}{r_{s3}^3} \\ -\frac{(1-\mu)z}{r_1^3} - \frac{\mu z}{r_2^3} - \frac{m_s(z-z_s)}{r_{s3}^3} \end{bmatrix}, \quad (9) \\
&\mathbf{r}_s = a_s \begin{bmatrix} \cos(\theta_s)\cos(\Omega) - \sin(\theta_s)\sin(\Omega)\cos(i) \\ \cos(\theta_s)\sin(\Omega) + \sin(\theta_s)\cos(\Omega)\cos(i) \\ \sin(\theta_s)\sin(i) \end{bmatrix}, \\
&\mathbf{a}_{srp} = \frac{p_{srp} c_R A_s}{m} \hat{\mathbf{r}}_{s3}, \quad \mathbf{r}_{s3} = \mathbf{r} - \mathbf{r}_s,
\end{aligned}
$$

where $\mathbf{X}_{opt} = [\Delta\mathbf{v}_0, \ldots, \Delta\mathbf{v}_k]$ represents the $k+1$ maneuvers, with $\Delta\mathbf{v}_0$ as the initial correction and $\Delta\mathbf{v}_k$ as the velocity difference at the final position. The term $\mathbf{e}_f$ is the final state error, and H penalizes deviations from the target state. The state vector $\mathbf{X}(t)$ consists of position $\mathbf{r} = (x, y, z)$ and velocity $\mathbf{v} = (\dot{x}, \dot{y}, \dot{z})$. μ, r_1, and r_2 follows the same definition as in the CR3BP formulation. The additional gravitational perturbation from the Sun is modeled using m_s and the relative position $\mathbf{r}_{s3} = \mathbf{r} - \mathbf{r}_s$, with $\mathbf{r}_s$ parameterized by a_s, i, Ω, and θ_s. The solar radiation pressure acceleration $\mathbf{a}_{srp}$ depends on p_{srp}, the reflectivity coefficient c_R, the spacecraft's cross-sectional area A_s, mass m, and unit direction $\hat{\mathbf{r}}_{s3}$.

The final velocity difference is added to the cost to achieve a smoother transition between two orbits and to extend the strategy for obtaining the annual cost for station-keeping. Additionally, a grid search is conducted to determine the optimal number of maneuvers for each orbit, with the results summarized in Table III. Station-keeping for orbits 1 and 2 would require numerous low-cost Δv impulses, some as small as 100 micrometers per second, which would be achieved using electric propulsion. Orbits 3 and 4 would use chemical propulsion, with a minimum Δv of around 3 centimeters per second. In this scenario, the refueler is assumed to carry both types of fuel. Furthermore, the reward is calculated based on the percentage of fuel rather than the absolute mass, to avoid bias toward refueling spacecraft with larger fuel capacities.

III. Numerical results

We present the results of our simulation, demonstrating the efficiency of the proposed refueling strategy for spacecraft in various cislunar orbits. The numerical analysis optimizes the refueling strategy over a twenty-year operational period. The scenario considers four mini-satellites, denoted as S/C$_i$ for $i \in \{1,2,3,4\}$, each with a dry mass of 500 kg (total mass without fuel). The fuel distribution is optimized to ensure efficient refueling operations. After testing different allocations, the fuel masses are set as follows: 80 kg for S/C$_1$ and S/C$_3$, 70 kg for S/C$_2$, and 55 kg for S/C$_4$.

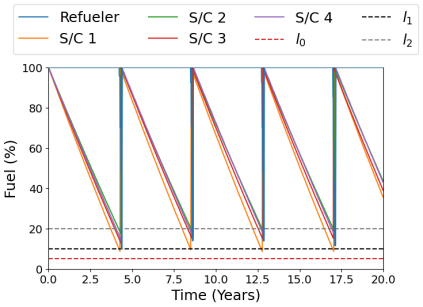
A dedicated refueling spacecraft, with a dry mass of 100 kg and an onboard fuel capacity of 700 kg, serves as the primary resource for sustaining mini-satellite operations. The refueling station, located in a DRO orbit, is assumed to have an unlimited fuel supply. The refueler employs high-performance chemical propulsion with a specific impulse (I_{sp}) of 315 seconds. Spacecraft 1 and 2 utilize electric propulsion with a specific impulse of 2000 seconds, while Spacecraft 3 and 4 use chemical propulsion with a specific impulse of 300 seconds. The fuel levels of the different spacecraft are defined as follows: l_0 is set at 5%, l_1 is at 10%, and l_2 at 20% of the associated satellite total fuel mass.

The strategy obtained by solving the MDP is compared with a greedy approach. This heuristic greedy strategy is defined using the following set of rules:

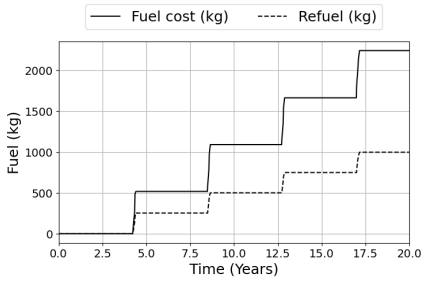
- When refueling begins, the refueler prioritizes the spacecraft with the lowest fuel level and continues refueling in ascending order until the process is complete
- If the refueler's fuel level drops below l_1, it must return to the refueling station.

The numerical results of the refueling simulations are shown in Fig. 5 and Fig. 6, which highlight the evolution of the fuel levels, the cumulative fuel usage, and refueling mass, and the refueling transfers over the twenty-year operational period. The transfers are shown directly between the system orbits without explicitly displaying the intermediate phasing orbits. However, the fuel and time costs for phasing maneuvers have been incorporated into the overall transfer cost and duration.

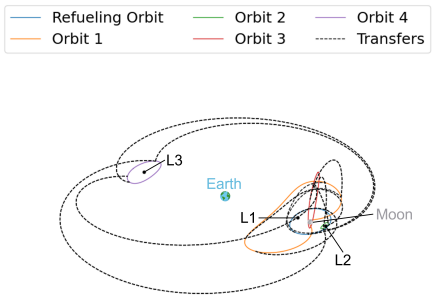
Fig. 5b shows the fuel cost and the refueled mass over the mission's duration using the MDP-based strategy. Compared to the greedy approach, this solution optimizes both the timing and trajectories of refueling maneuvers to ensure a more efficient fuel distribution while minimizing unnecessary trips between the refueling station and the satellites. In contrast, the greedy strategy in Fig. 6b leads to higher fuel consumption, as it results in more frequent refueling maneuvers and greater overall fuel expenditure due to a less coordinated approach. The lack of long-term planning forces the refueler to react to low fuel levels rather than scheduling refueling at optimal times, increasing operational demands and reducing efficiency. Ultimately, for the same total mass refueled, the MDP-based strategy consumed only 2200 kg of fuel, a 36% reduction compared to the greedy strategy. Overall, by considering long-term efficiency rather than



(a) Evolution of fuel levels

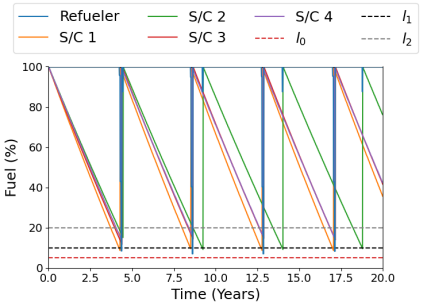


(b) Cumulative fuel consumption and total refueled mass

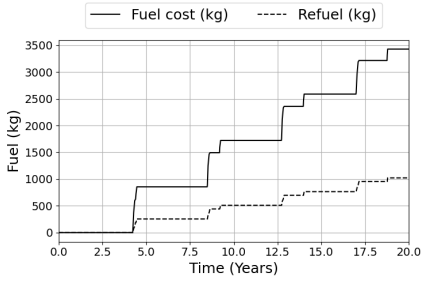


(c) Transfer trajectories for refueling maneuvers

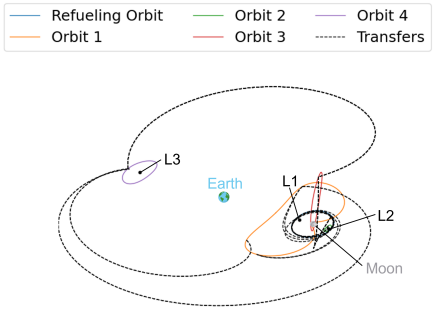
Fig. 5: Refueling performance for four spacecraft over 20 years with a strategy solution of the MDP. (a) Evolution of fuel levels, (b) cumulative fuel consumption and total refueled mass, (c) transfer trajectories for refueling maneuvers.



(a) Evolution of fuel levels



(b) Cumulative fuel consumption and total refueled mass



(c) Transfer trajectories for refueling maneuvers

Fig. 6: Refueling performance for four spacecraft over 20 years with a greedy strategy. (a) Evolution of fuel levels (b) cumulative fuel consumption and total refueled mass, (c) transfer trajectories for refueling maneuvers.

immediate refueling needs, the MDP strategy ensures that all spacecraft maintain operational capability while minimizing resource expenditure.

IV. Conclusion

The method developed enables efficient multi-spacecraft refueling operations in cislunar space. By integrating multi-fidelity models such as the two-body, circular restricted three-body, and bicircular restricted four-body models, we achieve precise and cost-effective refueling strategies. Simulations show a 36% reduction in fuel usage with the MDP-based strategy compared to a greedy refueling. This approach significantly improved efficiency by reducing refueling frequency, minimizing returns to the refueling orbit, and optimizing fuel costs through transfer sequencing. Future work will explore low-cost transfers for the refueler, the use of multiple refuelers to improve overall efficiency, and the optimization of refueling station locations to reduce fuel consumption further and enhance mission longevity.

References

[1] S. Do, R. Shishko, D. Antonelli, T. Cichan, R. Collom, P. Conrad, V. Coverstone, R. M. Davis, C. M. Edwards, M. Fuller *et al.*, "Logistics is a key enabler of sustainable human missions to mars," *Bulletin of the American Astronomical Society*, vol. 2, 2019.

[2] T. Taylor, W. Kistler, and R. Citron, "On-orbit fuel depot deployment, management and evolution focused on cost reduction," in *AIAA SPACE 2009 Conference & Exposition*, 2009, p. 6757.

[3] M. J. Holzinger, C. C. Chow, and P. Garretson, "A Primer on Cislunar Space."

[4] J. A. Vedda, "Cislunar development: What to build and why," *The Aerospace Corporation*, 2018.

[5] D. J. Tiffin and P. D. Friz, "Parametric cost analysis of refueling options in cislunar space," in *AIAA ASCEND*, 2021.

[6] JPL. Three-body periodic orbits. [Online]. Available: https://ssd.jpl.nasa.gov/tools/periodic_orbits.html

[7] M. Dominguez and K. C. Howell, "Preliminary design strategy for long-term loitering orbits in cislunar space," in *AAS/AIAA Astrodynamics Specialist Conference*, 2024.

[8] B. B. Jagannatha and K. Ho, "Event-driven network model for space mission optimization with high-thrust and low-thrust spacecraft," *Journal of Spacecraft and Rockets*, vol. 57, no. 3, pp. 446–463, 2020.

[9] C. Baier and J.-P. Katoen, *Principles of model checking*. MIT press, 2008.

[10] D. Vallado and W. McClain, *Fundamentals of Astrodynamics and Applications*, ser. Fundamentals of Astrodynamics and Applications. Microcosm Press, 2013.

[11] B. K. Sharma, "The Criteria for Reducing Centrally Restricted Three-Body Problem to Two-Body Problem," *International Journal of Astronomy and Astrophysics*, vol. 14, Mar. 2024.

[12] K. Boudad, "Disposal dynamics from the vicinity of near rectilinear halo orbits in the earth-moon-sun system," Ph.D. dissertation, Purdue University, 12 2018.

[13] M. Ghorbani and N. Assadian, "Optimal station-keeping near earth–moon collinear libration points using continuous and impulsive maneuvers," *Advances in Space Research*, vol. 52, no. 12, pp. 2067–2079, 2013.