Clustering Spacecraft Trajectories Generated from a Local Initial State Set

Renee L. Spear* and Natasha Bosanac†
University of Colorado Boulder, Boulder, Colorado, 80303

I. Introduction

A challenge that is emerging in space situational awareness and space traffic management is analyzing the possible future motions of an object in the presence of uncertainty within cislunar space. In a chaotic system, uncertainty in the initial state may result in a variety of possible paths over sufficiently long time intervals. The level of uncertainty associated with a state estimate may also increase when off-nominal conditions are present, the objects are difficult to observe and track over time, and when characterizing an unknown object [1]. As the initial state estimate, level of uncertainty, and model parameters all vary, the array of possible future motions can evolve in geometry and itinerary as the paths pass through distinct regions of the Earth-Moon system. In these motivating examples, it may be valuable to extract a clear, digestible summary of the distinct motion types that originate from a local region of the phase space.

Sets of nonlinear trajectories may be challenging to manually categorize in the absence of an analytical solution or require automated analysis when reducing the dependency on a human-in-the-loop. In these cases, data mining techniques may be useful for automatically extracting patterns and other information [2]. One technique, clustering, enables extraction of a set of clusters that group similar data and separate dissimilar data [2]. These clusters can supply an automated and digestible summary of a larger dataset that aids subsequent analysis.

Trajectory clustering, specifically, focuses on discovering patterns or groups of possible motions for a moving object [3]. There are various approaches to summarizing the trajectories to be clustered, including 1) using a time series representation, which may offer a high-fidelity representation of a trajectory but suffer from the curse of dimensionality; or 2) constructing a model of the trajectory, which may reduce the dimension of the description but suffer from limited accuracy when applied to a wide variety of paths [3]. These finite-dimensional trajectory descriptions are then input to a selected clustering algorithm to extract groups of sufficiently similar paths.

Clustering has previously been used in astrodynamics. Early examples include Hadjighasem et al. using spectral clustering for detecting Lagrangian vortices within nonlinear dynamical systems [4]; Nakhjiri and Villac using k-means clustering for locating bounded motions near a distant retrograde orbit on a Poincaré map [5]; and Foslien et al.

*Graduate Researcher, Ann and H.J. Smead Department of Aerospace Engineering Sciences, 3775 Discovery Dr., Boulder, CO, 80303. AIAA Student Member. Corresponding author: renee.spear@colorado.edu

†Assistant Professor, Ann and H.J. Smead Department of Aerospace Engineering Sciences, 3775 Discovery Dr., Boulder, Boulder, CO, 80303. AIAA Senior Member.

An earlier version of this Note was presented as paper AAS 23-110 at the AAS/AIAA Astrodynamics Specialist Conference in August 2023 at Big Sky, MT.

using data mining to identify anomalous conditions on the International Space Station's gimbal system [6]. More recently, Bosanac as well as Bonasera and Bosanac have developed a distributed clustering framework to summarize the geometries of a wide array of trajectories in a multi-body system [7–9]. Smith and Bosanac also used clustering to summarize arcs along families of periodic orbits and hyperbolic invariant manifolds that supply building blocks for trajectory design in a multi-body system [10]. Finally, Miceli et al. used clustering to extract low lunar frozen orbits from trajectories generated in a high-fidelity lunar gravity model [11]. These examples have demonstrated the value of clustering for automatically extracting information from datasets in astrodynamics.

This Note presents a clustering approach to categorizing possible spacecraft trajectories generated from a local point cloud of initial states by their geometric and temporal variations. In this proof of concept, initial states are propagated forward in time in the Earth-Moon circular restricted three-body problem (CR3BP) to generate the associated continuous trajectories. Each trajectory is then discretized into a set of states that are evenly distributed as a function of arclength. Finite-dimensional feature vectors are constructed to summarize each continuous trajectory using the velocity direction at each sample and elapsed time between samples. Then, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) [12, 13] is used to cluster these feature vectors. Finally, these clusters are updated with additional information to increase their accuracy and ensure less prevalent geometries are captured, if they exist. The result is a set of clusters that supply a digestible summary of the array of possible motion types emanating from a local point cloud of initial states. This process is applied to trajectory sets that begin in the local neighborhood of reference states along periodic orbits near L_1 and a distant prograde orbit (DPO).

This Note offers an original contribution to characterizing trajectories that emanate from a local region of the phase space but exhibit diverse paths through a chaotic system. This challenge has been encountered in propagating uncertainty along spacecraft trajectories. Many uncertainty propagation analyses focus on trajectories that closely resemble a single nominal path. However, recent approaches have addressed calculating probability distributions that describe non-Gaussian variations around a reference path or approximating distinct modes governing the evolution of a diverse set of trajectories using Gaussian mixture models [14], radial basis functions [15], and the probabilistic admissible region [16]. In contrast, the method presented in this Note focuses on using clustering to group a diverse set of paths that begin from a local region of the phase space by their geometry. This approach can discover geometrically distinct groups without assuming a function approximation form, incorporate trajectory characteristics, and avoid propagating additional parameters along the trajectories. Furthermore, unlike previous applications of clustering to spacecraft trajectories [8, 9, 11], the local proximity of the initial conditions of trajectories generated in this Note reduces both the geometric diversity across the dataset and separation between geometrically different trajectories. These characteristics present an interesting new application for using clustering to discover distinct motion types. In contrast to these works, this Note also uses a slightly different trajectory summarization and description approach as well as an additional resampling step.

II. Dynamical Model

The CR3BP is used to approximate the motion of a spacecraft of assumed negligible mass under the gravitational influence of two massive bodies, labeled primaries [17]. In this Note, the primaries are the Earth and the Moon which are modeled as point masses with constant mass and travel on circular orbits about their barycenter [17]. Length, mass, and time quantities are nondimensionalized to set the distance between the Earth and Moon, the total mass of the primaries, and the mean motion of the primary system to unity. A rotating reference frame is defined with the origin at the Earth-Moon barycenter and axes $\hat{x}\hat{y}\hat{z}$: $\hat{x}$ points from the Earth to the Moon, $\hat{z}$ is aligned with the orbital angular momentum vector of the primary system, and $\hat{y}$ completes the orthogonal, right-handed triad [17]. Throughout this Note, all trajectories are visualized in this rotating frame. Given a nondimensional state vector that is defined in the rotating frame as $\boldsymbol{x} = [x, y, z, \dot{x}, \dot{y}, \dot{z}]^T$, the equations of motion for a spacecraft in the CR3BP are written as

$$\ddot{x} - 2\dot{y} = \frac{\partial U}{\partial x}, \qquad \ddot{y} + 2\dot{x} = \frac{\partial U}{\partial y}, \qquad \ddot{z} = \frac{\partial U}{\partial z} \qquad \text{where} \qquad U = \frac{\left(x^2 + y^2\right)}{2} + \frac{(1-\mu)}{r_1} + \frac{\mu}{r_2} \tag{1}$$

In these equations, $\mu \approx 1.21505842695 \times 10^{-2}$ is the ratio of the Moon's mass to the total system mass, $r_1 = \sqrt{(x+\mu)^2 + y^2 + z^2}$, and $r_2 = \sqrt{(x-1+\mu)^2 + y^2 + z^2}$. In the rotating frame, the Jacobi constant is equal to [17]

$$C_J = \left(x^2 + y^2\right) + \frac{2\,(1-\mu)}{r_1} + \frac{2\mu}{r_2} - \dot{x}^2 - \dot{y}^2 - \dot{z}^2 \tag{2}$$

At a single value of this quantity, a wide variety of motions may exist. In particular, the space community commonly leverages five libration points, L_i where $i \in [1, 5]$, and families of periodic orbits for trajectory design [17, 18]. Numerical approaches to computing periodic orbit families appear throughout the astrodynamics literature, such as in Refs. [18] and [19]. In this Note, differential correction and continuation schemes are used to compute periodic orbits.

III. Clustering

Clustering is a data mining technique used to group similar members of a dataset and separate dissimilar members [3]. Consider N members of a dataset to be clustered. Each member is described by an M-dimensional feature vector, $\boldsymbol{f}_i$ for $i \in [1, N]$, that is tailored to the desired application. Then, each pair of members is compared by calculating the distance between their feature vectors using a selected distance measure [2]. These distances are used by a clustering algorithm, along with any governing parameters, to construct groups of sufficiently similar members of the dataset. Various clustering algorithms exist, with the most common approaches classified as partitioning, model-based, hierarchical, or density-based methods [2]. Selection of a suitable algorithm is influenced by the characteristics of the data and the application; this decision, along with the values of any governing parameters, may also influence the recovered clusters.

This Note uses HDBSCAN, a clustering algorithm that was developed by Campello et al. as a hierarchical extension of the well-known Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [12]. By combining density-based and hierarchical clustering concepts, HDBSCAN supports the identification of clusters as dense groupings of data in the selected M-dimensional feature space while also allowing clusters to possess distinct shapes and densities [12, 13]. Furthermore, this approach does not require a priori knowledge of the number of clusters [12, 13]. Because of these properties, this algorithm has previously been used to group spacecraft trajectories by their geometry in the work performed by Bosanac [7, 9], Bonasera and Bosanac [8], and Miceli et al. [11].

HDBSCAN first assesses the similarity between the N members of a dataset using density information. A core distance d_{core} is calculated for each member as the distance to its $N_{minCore}$-nearest neighbor. The core distances of members a and b are used to calculate their mutual reachability distance as $d_{mreach}(a, b) = max\{d_{core}(a), d_{core}(b), d(a, b)\}$ where $d(a, b)$ indicates a selected distance measure [12]. This transformation to a mutual reachability distance space further separates members within high-density regions from those in sparse regions. Furthermore, $N_{minCore}$ serves as a governing parameter that defines the size of the neighborhood used to capture density information.

HDBSCAN then selects groupings from a cluster hierarchy and labels each member of the dataset by a cluster assignment or as noise. Then, a minimum spanning tree is constructed to summarize a weighted graph representation of the mutual reachability distances between the members of the dataset [12]. An agglomerative approach is applied to this minimum spanning tree to create a hierarchy of possible clusters as a function of the mutual reachability distance. The final clustering results are selected from this hierarchy to construct groups that contain at least $N_{minClust}$ members and maximize the cluster stability or persistence [12]. Malzer and Baum present a modification of the HDBSCAN algorithm that merges clusters separated by a distance below a threshold ϵ [20]; this modification mitigates the possibility of recovering an excessive number of small clusters that correspond to locally dense regions within a large group of sufficiently similar members. Any members not grouped are classified as noise points and may represent outliers in the data or insufficiently sampled regions. HDBSCAN is accessed in Python using the *hdbscan* package which requires a computational effort of $O(N \log(N))$ and is predominantly governed by $N_{minCore}$, $N_{minClust}$, and ϵ [21].

IV. Data-Driven Categorization Process

This section presents a clustering-based approach for automatically categorizing a diverse array of trajectories generated from a set of initial states that begin near a reference state. Each of the generated clusters supplies a 'motion type' that captures a set of trajectories with similar geometric and temporal variations. As a result, the collection of clusters supplies a digestible summary of the set of trajectories. The data-driven categorization process consists of five steps that are depicted in the flowchart in Fig. 1 and described in detail within this section. Furthermore, each step of the technical approach is demonstrated for planar trajectories generated from a set of initial states in the vicinity of a reference state along an L_1 Lyapunov orbit in the Earth-Moon CR3BP. The computational times for each step are determined using a computer with an AMD Ryzen 5 5600G with Radeon Graphics 4.20 GHz processor.

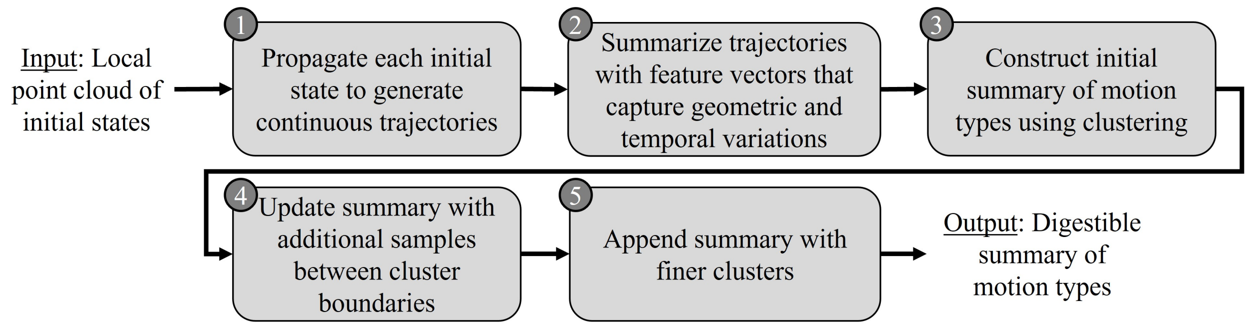

Fig. 1 High-level overview of the data-driven categorization process.

A. Generating Input Data

The data-driven categorization process takes an input of a set of initial states that lie in the local neighborhood of a reference state. Throughout this Note, this set is labeled as a local point cloud of initial states, encompassing a hyperellipsoid in six-dimensional space that is centered on the selected reference state. This local point cloud could physically represent a snapshot of the state space where an object is reasonably expected to exist at a selected epoch within, for example, 1) a 3-σ uncertainty region around a state estimate or 2) a pre-specified corridor for targeting a reference state. These motivating examples often involve associating probabilities to each initial state. However, this Note relies on sampling initial conditions within the boundary of the local point cloud to identify the dominant geometries of the associated trajectories regardless of their likelihood of occurrence. Future work that applies this data-driven categorization process to trajectory prediction problems could incorporate probabilities during sampling, categorization, or after generating a digestible summary of possible motions.

The scheme used to sample initial states within the local point cloud balances supplying a sufficiently dense set of samples with reducing the total number of samples for computational feasibility. Although uniformly sampling each of the six dimensions of the phase space within a specified region around a reference state is conceptually straightforward, this approach was observed to suffer from sparse coverage when generating a reasonable number of samples [22]. To reduce the number of samples while sufficiently covering the local region of the phase space, position and velocity vectors are first independently sampled in this Note using a grid-based approach that supplies dense and uniform sampling only in their associated three-dimensional subspaces. Samples are defined using a square, uniformly-spaced grid and only those n_{IC} position vectors or n_{IC} velocity vectors that lie within the specified boundaries of the local point cloud are retained. Then, the pseudorandom number generator in MATLAB's *randperm* function is used to identify two distinct sequences of unique integers between 1 and n_{IC} [23]. The ith integer in each list identifies the specific position and velocity vectors that are combined to form the ith of n_{IC} six-dimensional state vectors.

To demonstrate the sampling approach used in this Note, consider a local point cloud of initial states that exist near a single reference state along an L_1 Lyapunov orbit with a period of 12.269 days and C_J = 3.1542 in the Earth-Moon CR3BP. This nondimensional reference state is specified in the rotating frame as $\boldsymbol{x}_{PO} = [0.816988444235, 0, 0, 0, 0.195756600373, 0]^T$, occurring at the leftmost crossing of the x-axis. Planar states that lie within 10.5 km in position components and 10.5 m/s in velocity components of the reference state are sampled using the method described in the previous paragraph. This local point cloud size corresponds to the 3-σ uncertainties estimated by Bradley et al. [24] for cislunar navigation using optical (angles only) measurements. Although Fedeler et al. [25] note that uncertainty levels may vary for different types of motions, the proof of concept in this Note uses the same local point cloud size for each of the selected scenarios. In this L_1 Lyapunov scenario, 1,006 states are sampled, which typically requires on the order of 10^{-2} seconds on the specified computer. Across this set of initial states, the Jacobi constant varies on the order of 10^{-3} nondimensional units.

B. Step 1: Propagate Continuous Trajectories

Each of the input initial states is propagated in a selected dynamical model for a specified duration. In this Note, all initial conditions are propagated forward in time for at least 17.3 days in the Earth-Moon CR3BP; this duration exceeds the periods of the selected orbits and allows distinct geometries to emerge, if they exist. However, propagation is terminated early upon impact with a spherical approximation of either of the primaries, defined using their equatorial radii of 1,738.0 km and 6,378.1363 km for the Moon and Earth, respectively [26]. Following this procedure, Fig. 2 displays a 200-member subset of the continuous trajectories generated from the 1,006 initial conditions used in this scenario with the L_1 Lyapunov orbit shown in blue. Generating the full set of continuous trajectories, $\mathcal{T}$, typically requires on the order of 0.1 seconds in C++ on the specified computer. These initial conditions occur near the red circle with the Moon and Earth appearing as gray circles scaled to possess the radius of the respective body, and red diamonds locating L_1 and L_2. Figure 2 demonstrates that paths either impact the Moon, remain in the vicinity of the Moon, or depart through one of the L_1 or L_2 gateways, producing a variety of distinct geometries.

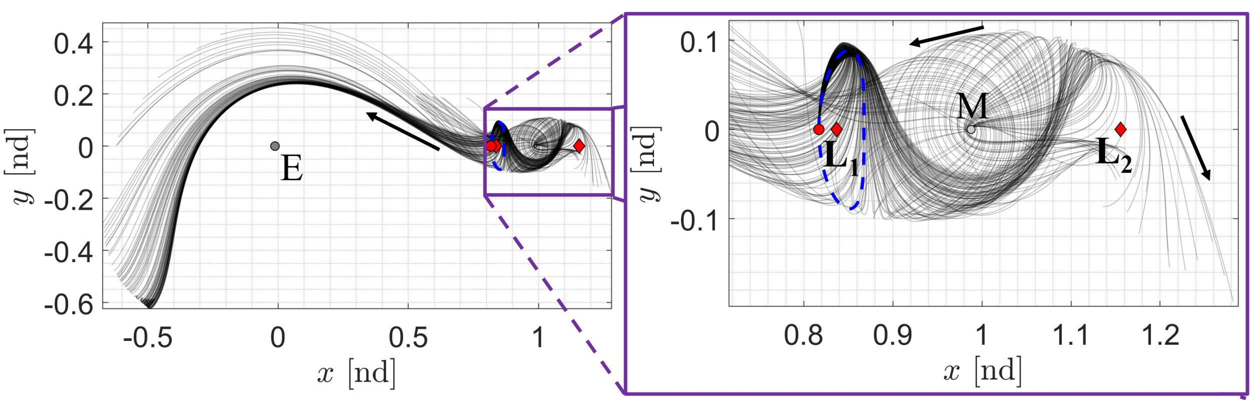

Fig. 2 A 200-member subset of planar trajectories propagated for 17.3 days from the sampled initial conditions in the Earth-Moon CR3BP.

C. Step 2: Summarize Trajectories

To enable categorization via clustering, each continuous trajectory is discretized into a sequence of p states at equal intervals in arclength. When numerically integrating a state from t_1 to t_2, the arclength, d_{arcL}, is defined as the distance traversed along the trajectory [27]. In this Note, this quantity is calculated along a trajectory in the rotating frame as

$$d_{arcL} = \int_{t_1}^{t_2} |\boldsymbol{v}|\, dt \tag{3}$$

where $\boldsymbol{v} = [\dot{x}, \dot{y}, \dot{z}]^T$ is the nondimensional velocity vector. Then, p states, including the final state but excluding the initial state, are sampled at intervals of d_{arcL}/p along the arclength of the trajectory. The initial state is excluded because all initial states lie within a local neighborhood and, therefore, do not meaningfully contribute to identifying geometric differences between trajectories. Although there are various possible approaches for discretizing a continuous path, this Note employs an equal arclength distribution to 1) limit the influence of variations in the speed along a trajectory, and 2) ensure sample locations smoothly vary as the geometry gradually evolves. However, a limitation of this approach is that the samples may not be placed in geometrically meaningful locations, such as apses, or sufficiently capture smaller revolutions along a trajectory that traverses a large distance when p is too small.

The number of states, p, used to discretize each trajectory is selected heuristically using a curve-based approach. The curvature, $\kappa(\boldsymbol{x})$, reflects the deviation from a straight line at a state $\boldsymbol{x}$ along a trajectory and is equal to [27]

$$\kappa(\boldsymbol{x}) = \frac{\sqrt{(\ddot{z}\dot{y} - \ddot{y}\dot{z})^2 + (\ddot{x}\dot{z} - \ddot{z}\dot{x})^2 + (\ddot{y}\dot{x} - \ddot{x}\dot{y})^2}}{\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right)^{3/2}} \tag{4}$$

Local maxima in the curvature supply geometrically meaningful locations along a trajectory and typically occur near periapses and apoapses measured relative to primaries or equilibrium points. However, unlike apses, maxima in the curvature do not require specification of a reference point as one or more trajectories exhibit turning points in distinct regions of the Earth-Moon system. To identify the number of local maxima in the curvature along a single trajectory, p_{traj}, the curvature is calculated at each state along the trajectory in the rotating frame via Eq. 4 and then the local peaks in the time history of $\kappa(\boldsymbol{x})$ are identified. The largest number of local maxima in the curvature along each trajectory in the dataset defines the value of p_{max}. Then, each trajectory in the dataset is sampled using $p = 2(p_{max} + 1)$ states that are equally-distributed in arclength. This value of p is selected to ensure there are at least twice as many samples as the number of significant geometric locations, including the final state, along the trajectory. As a result, this value of p has been observed to sufficiently summarize all trajectories in the datasets examined in this Note. Furthermore, each trajectory in the dataset is summarized by the same number of samples, reducing the complexity of clustering.

To demonstrate the trajectory discretization process, consider a single, planar trajectory generated in Step 1. Figure 3a) displays this trajectory along with blue stars at the locations of maximum curvature and a black star at the final state. Repeating this calculation for each of the 1,006 trajectories generated for 17.3 days in Step 1 produces $p_{max} = 7$, consistent with the number of curvature maxima that occur along the trajectory in Fig. 3a). This trajectory is then discretized into $p = 16$ states that are spaced equally in arclength. Figure 3b) shows the resulting discretization with the sampled states, indicated by black stars, sufficiently capturing the shape of the path within the configuration space.

The p sampled states are used to form two finite-dimensional feature vectors that, together, capture the spatiotemporal evolution of a trajectory. Although position-based feature vectors are a viable option for encoding the spatial evolution of a trajectory, different regions of a multi-body system possess distinct sensitivities to the same distance between two sampled position vectors. This challenge can result in lower quality clusters generated from sequences of position vectors, including excessive grouping, incorrect grouping, or excessive differentiation. Furthermore, it is difficult to intuitively and automatically select suitable tolerances for two trajectories to be considered sufficiently similar. To address this challenge, Bosanac has demonstrated using velocity unit vectors to cluster spacecraft trajectories by their shape [28]. Applying this concept to this Note, all trajectories begin from a single local neighborhood in the phase space. Therefore, two trajectories with a similar shape and time evolution are expected to follow similar paths through the configuration space. Accordingly, the following two separate feature vectors are used in this Note: 1) $\boldsymbol{f}_{\hat{v}}$ captures the shape of the path via the velocity direction at sampled states and 2) $\boldsymbol{f}_{\Delta\tau}$ captures the temporal evolution via the elapsed time between subsequent states. Mathematically, these feature vectors are calculated as

$$\boldsymbol{f}_{\hat{v}} = \left[\hat{\dot{x}}_1, \hat{\dot{y}}_1, \hat{\dot{z}}_1, \cdots, \hat{\dot{x}}_p, \hat{\dot{y}}_p, \hat{\dot{z}}_p\right] \qquad \boldsymbol{f}_{\Delta\tau} = \left[\Delta\tau_1, \cdots, \Delta\tau_p\right] \tag{5}$$

where $\hat{\dot{x}}_i, \hat{\dot{y}}_i, \hat{\dot{z}}_i$ are the components of the velocity unit vector at sampled state i in the rotating frame. In addition, $\Delta\tau_i$ is the nondimensional time elapsed from the previous state, $\boldsymbol{x}_{i-1}$, for $i > 1$ or the initial time, $\tau_0 = 0$, when $i = 1$, to the current state, $\boldsymbol{x}_i$. Through these definitions, $\boldsymbol{f}_{\hat{v}}$ is a $3p$-dimensional vector for spatial motion whereas $\boldsymbol{f}_{\Delta\tau}$ is a p-dimensional vector. For planar trajectories, the z-component of the velocity unit vector at each state is removed from $\boldsymbol{f}_{\hat{v}}$ to avoid including redundant information. In the selected scenario, summarizing the full set of trajectories

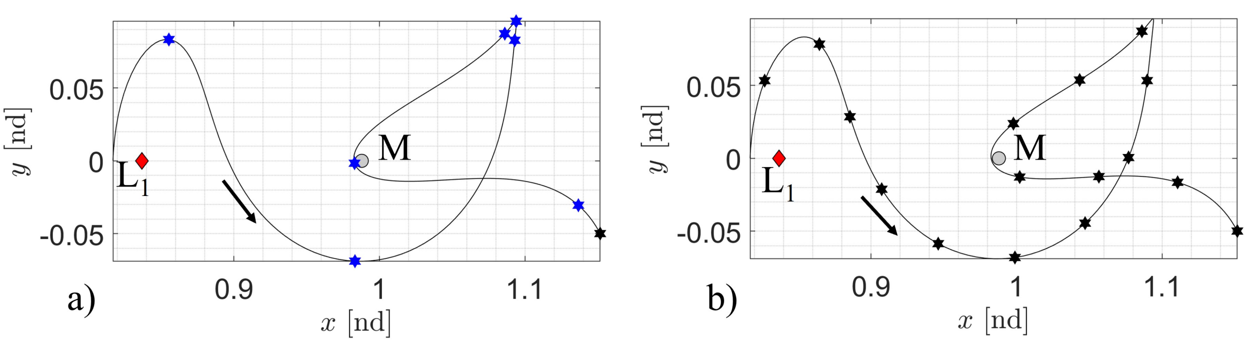

Fig. 3 a) Locations of local maximum curvature along a trajectory. b) The locations of p states spaced equally in arclength along the trajectory.

typically requires around four seconds on the specified computer with $O(N)$ complexity for each step. The process for summarizing trajectories is outlined in the flowchart provided in Fig. 4.

D. Step 3: Construct Initial Summary of Motion Types

With a selected set of input parameters, HDBSCAN is used to automatically group trajectories by their spatiotemporal variations, producing an initial summary of motion types. This clustering process is summarized in the flowchart provided in Fig. 5. First, trajectories are clustered using their velocity direction feature vectors, $\boldsymbol{f}_{\hat{v}}$, to discover groups of geometrically similar trajectories. Although the trajectories are generated from a set of initial states that lie within a local region of the phase space, their geometric variations over sufficiently long time horizons are expected to result in dense groupings of the data within the higher-dimensional feature space associated with $\boldsymbol{f}_{\hat{v}}$. Each dense grouping corresponds to a cluster that should be identified by HDBSCAN. Then, each cluster is refined by clustering only its members via their time-based feature vectors $\boldsymbol{f}_{\Delta\tau}$. This cluster refinement step supports outlier removal and allows groups to be further separated, as needed, based on additional temporal variations. After this initial two-step clustering process, each trajectory is either assigned to a group or labeled as noise.

Although $\boldsymbol{f}_{\hat{v}}$ and $\boldsymbol{f}_{\Delta\tau}$ could be weighted and combined to construct a single $4p$-dimensional feature vector, simultaneously clustering a dataset with distinct features can result in a higher likelihood of inaccurate groupings. Specifically, the same difference between two weighted but distinct feature types could result in two trajectories being assigned to the same cluster despite substantial differences in their geometric and/or temporal variations. However, independently using the two feature vectors has been observed in this work to improve the cluster accuracy while also eliminating the need to strategically weight the distinct features.

The parameters governing HDBSCAN are manually defined to be consistent across all scenarios in this Note. However, automated and adaptive parameter selection is a valuable avenue of future work. Manual selection of the minimum cluster size and neighborhood size begins with setting $N_{minCore} - 1 = N_{minClust}$ to measure density across the same-sized neighborhood as the minimum acceptable cluster size. Then, a small value of $N_{minClust}$ is selected to prioritize local density estimates and variations between nearby members of the data. In addition, the Euclidean distance is selected to assess the difference between two feature vectors to support fast clustering. Finally, the minimum

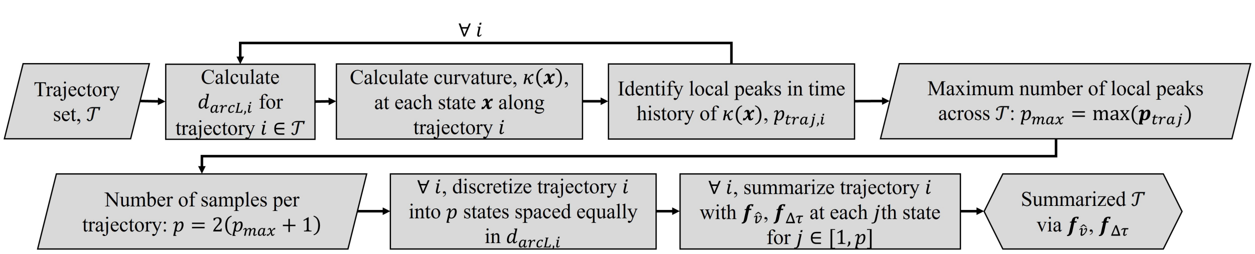

Fig. 4 Overview of the trajectory summarization process in Step 2.

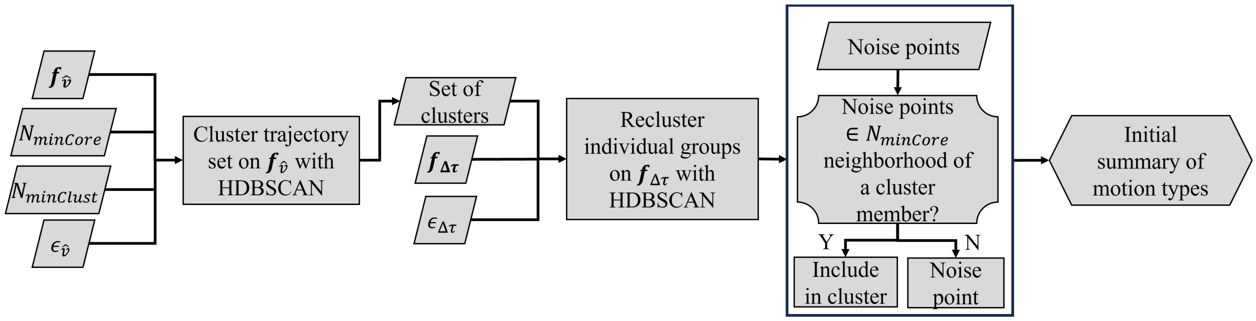

Fig. 5 Overview of the process for generating an initial motion type summary in Step 3.

threshold for cluster differentiation in the velocity-based feature space is heuristically selected as $\epsilon_{\hat{v}} = 2\sqrt{p}\sin\left(\frac{\alpha}{2}\right)$. This heuristic, as presented by Bosanac, reflects the Euclidean distance between a set of p unit vectors separated by an angle equal to α [28]. For cluster refinement in the time-based feature space, $\epsilon_{\Delta\tau} = N_{minCore}\max(k, \epsilon_{thresh})$ where k is the maximum core distance from each trajectory to its nearest neighbor and ϵ_{thresh} is a user-selected small minimum threshold. Similar to the heuristic presented by Bosanac, this quantity estimates a reasonable $N_{minCore}$-neighborhood radius that adapts to the trajectories within the cluster to be refined and is not biased by significant outliers [28].

After a set of clusters has been generated, border point association is performed to label noise points that lie in the local neighborhood of a cluster member. This post-processing step is suggested by Campello et al. in an early paper on HDBSCAN [13]. Specifically, any noise points in the trajectory set, $\mathcal{T}$, that lie within the $N_{minCore}$-neighborhood of a labeled member in the mutual reachability distance space are considered border points of their closest cluster [13].

Applying this clustering process to the trajectories generated from the local point cloud of initial states near the selected reference state along a 12.269-day L_1 Lyapunov orbit results in successful separation of spatiotemporally distinct trajectories. The 1,006 trajectories are grouped into 13 clusters using $N_{minCore} = 4$ and $N_{minClust} = 5$ for clustering in both $f_{\hat{v}}$ and $f_{\Delta\tau}$, $\epsilon_{\hat{v}}$ is calculated using $\alpha = 5$ degrees, and $\epsilon_{\Delta\tau}$ is calculated using $\epsilon_{thresh} = 10^{-3}\sqrt{p}$. On the specified computer, this step requires on the order of one second to complete. The initial summary is presented in Fig. 6 with each subfigure displaying one cluster of trajectories in the Earth-Moon rotating frame. The L_1 Lyapunov orbit is depicted with a black dashed line in cluster 11. The horizontal and vertical axes represent the x- and y-position coordinates, respectively, in the Earth-Moon rotating frame in nondimensional units. In addition, the final state of each trajectory is denoted with a blue dot whereas color indicates the elapsed time along each trajectory. Black arrows denote the direction of motion. Twelve trajectories are categorized by HDBSCAN as noise points and are not depicted in this summary. This noise rate, approximately 1.2%, is considered acceptable because these noise points either 1) resemble existing clusters but do not lie within the $N_{minCore}$-neighborhood of cluster members, or 2) do not exist near enough trajectories with similar variations in shape and time to create a separate cluster of at least $N_{minClust}$ members.

Figure 6 reveals that each cluster contains trajectories that are geometrically similar. For instance, clusters 1-4

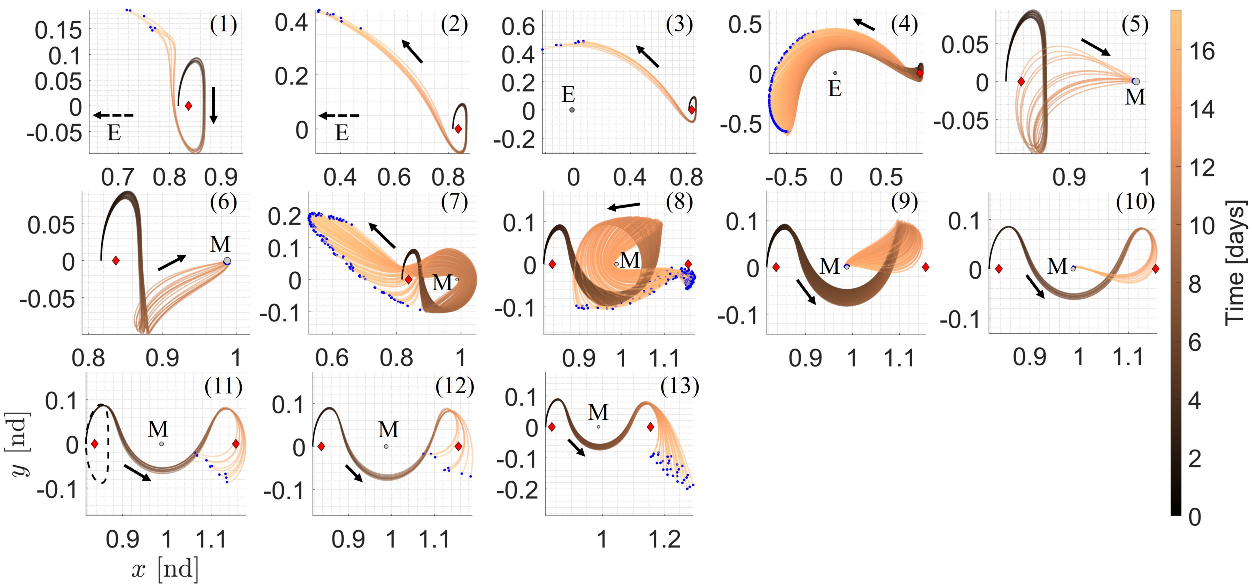

Fig. 6 Initial summary of trajectories generated from a local point cloud of initial states along the selected L_1 Lyapunov orbit in the Earth-Moon CR3BP.

capture trajectories that briefly resemble three-quarters of an L_1 Lyapunov orbit before departing for the vicinity of the Earth. Clusters 7-8, however, contain trajectories that resemble half an L_1 Lyapunov orbit before revolving around the Moon once with a low perilune. These trajectories are separated into two clusters because there is a distinct split in the velocity direction along the trajectories.

E. Step 4: Update Summary with Additional Samples

To increase sampling near cluster boundaries or in sparse regions between clusters, additional initial states are sampled from within the local point cloud and used to update the summary produced in the previous step. To contextualize the formulation of this process, spatiotemporal variations between the trajectories in distinct clusters from Step 3 are observed to be correlated to variations in the velocity components of the initial state. To demonstrate this observation, Fig. 7a) displays the velocity components of each initial condition colored by the assigned cluster with white diamonds locating points designated as noise by HDBSCAN. Examining the connection between the clustering results and the structure of the local neighborhood of the reference trajectory is an interesting avenue of future work. Nevertheless, in this example, the majority of noise points lie between cluster boundaries. Therefore, expanding the trajectory set with additional samples between these boundaries may aid in grouping noise points along with increasing the accuracy of cluster boundaries.

The inter-cluster sampling scheme is inspired by a Voronoi diagram which divides a g-dimensional space into cells such that each cell consists of one point and cell boundaries lie at the midpoint between two points [29]. In this Note, every velocity vector from the original local point cloud of initial states is treated as if it lies in its own cell in velocity space. This approach enables the sampling of additional initial velocities without relying on explicit boundary construction. Then, cluster labels for the trajectories generated from the velocity vectors of two initial states that lie in adjacent cells are compared. If the cluster labels are different, a new velocity vector is sampled at the midpoint between the original two velocity vectors. To complete the construction of a six-dimensional state vector, the position vector of this new initial state is recovered by placing a sample at the midpoint between the position components corresponding to the two original velocity samples. This process is summarized in the flowchart provided in Fig. 8. Inter-cluster sampling for the L_1 Lyapunov orbit scenario used throughout this section produces an additional 536 planar initial conditions which fall between cluster boundaries, shown in black in Fig. 7b).

The trajectories associated with the expanded set of initial conditions are clustered with HDBSCAN to construct a new summary. First, the new trajectories are summarized with the method presented in Step 2. Then, all trajectories are clustered using the same method and governing parameters presented in Step 3 followed by border point association. The result is an updated set of clusters and noise points.

This step sometimes results in some of the clusters from the initial summary being joined as two or more separated groups gain sufficient new samples between them to identify a continuous variation between members of the two groups. As an example, consider the expanded dataset generated from initial states near the leftmost x-axis crossing of the 12.269-day L_1 Lyapunov orbit in the Earth-Moon CR3BP. In this case, clusters 9-13 of Fig. 6 are joined to produce a single group of trajectories that briefly resemble the L_1 Lyapunov orbit for approximately one quarter of a revolution before impacting the Moon, performing a close approach to the Moon with a low perilune, or departing through the L_2 gateway. Although the trajectories terminate in different regions, their shape continuously varies within the cluster and over this time horizon. Alternative sampling schemes and feature vector definitions may enable HDBSCAN to separate

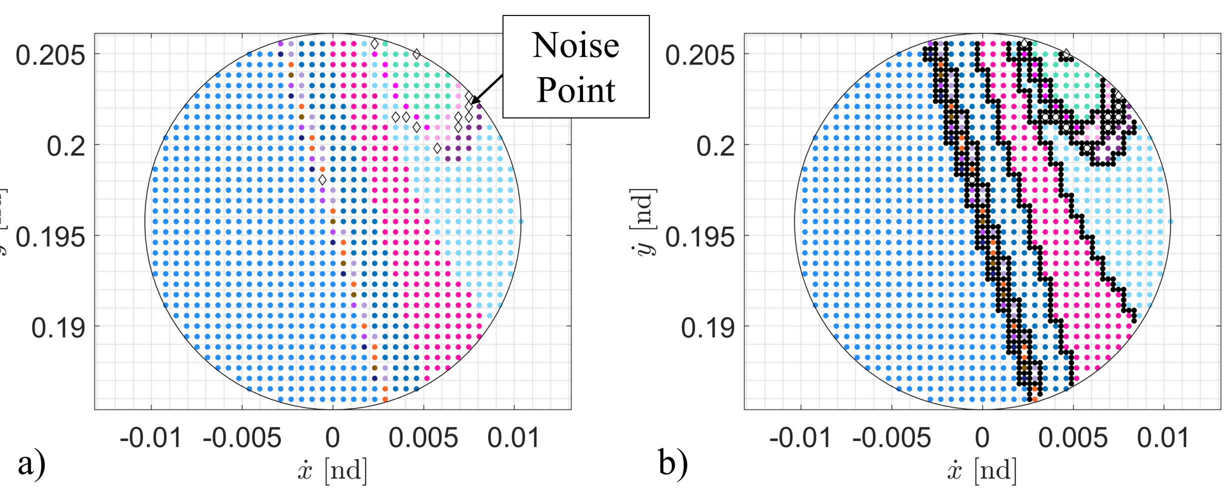

Fig. 7 Velocity components of initial conditions a) after initial clustering and b) with 536 new samples (black circles) after inter-cluster sampling.

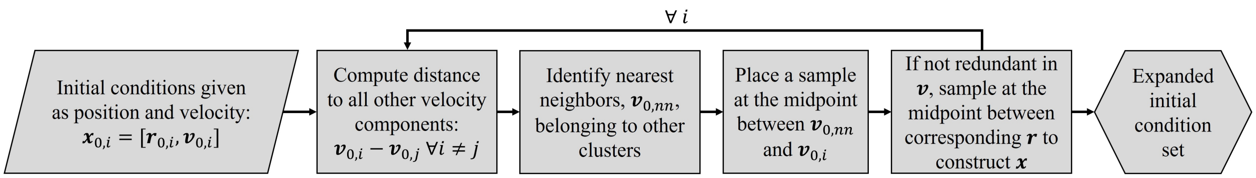

Fig. 8 Overview of the inter-cluster sampling process in Step 4.

this large cluster of trajectories as their distribution in the higher-dimensional feature vector space evolves, thereby producing a more detailed summary with a higher number of smaller clusters. For this scenario, this step requires around eight seconds on the specified computer with a computational complexity of $O(N^2)$.

F. Step 5: Append Summary with Finer Clusters

The final step of the data-driven categorization process involves grouping the remaining noise points to capture finer clusters that may consist of fewer members. Specifically, there may be some trajectories that are designated as noise, but are similar to fewer than $(N_{minClust} - 1)$ members. Accordingly, this final step leverages the clustering method presented in Step 3 to identify smaller groups of noise points that correspond to insufficiently sampled or low-density regions in the larger dataset. In this case, the HDBSCAN input parameters are manually selected with smaller values of $N_{minCore}$ and $N_{minClust}$ to construct locally-dense groups with $\epsilon_{\Delta\tau} = \epsilon_{\hat{v}} = 0$ due to the small number of trajectories in this dataset. Then, these new, smaller clusters are appended to the summary of the expanded dataset produced in Step 4. After this step, some trajectories may still be designated as noise.

Applying this final step to the remaining unlabeled trajectories after Step 4 in the 12.269-day L_1 Lyapunov orbit scenario in the Earth-Moon CR3BP allows the discovery of finer groupings from amongst the unlabeled trajectories. Selecting $N_{minCore} = 1$ and $N_{minClust} = 2$ results in eight small clusters that are appended to the cluster summary generated in Step 4, producing a total of 21 clusters; selecting $N_{minCore} = 1$ removes the notion of density when calculating the mutual reachability distance. This step produces a larger number of more detailed clusters from the reduced dataset. Using the same representation as Fig. 6, Fig. 9 displays the 21 motion types that capture the array of geometries exhibited by 1,542 trajectories as well as the remaining six trajectories that are designated noise. This final step in the data-driven process typically requires on the order of 10^{-2} seconds on the specified computer. The total runtime for the example used throughout this section is 25.25 seconds.

Analyzing Fig. 9 reveals that each cluster successfully captures trajectories with a similar geometry whereas distinct clusters separate trajectories with distinct geometries. For instance, clusters 20-21 produced in Step 5 are noticeably different from the trajectories in the larger clusters. Other small clusters added in this step resemble existing larger clusters but either possess a sufficiently different shape or temporal evolution, or exist in a more localized region of high density. The remaining six trajectories that are designated as noise resemble the members of clusters 5, 8, 14,

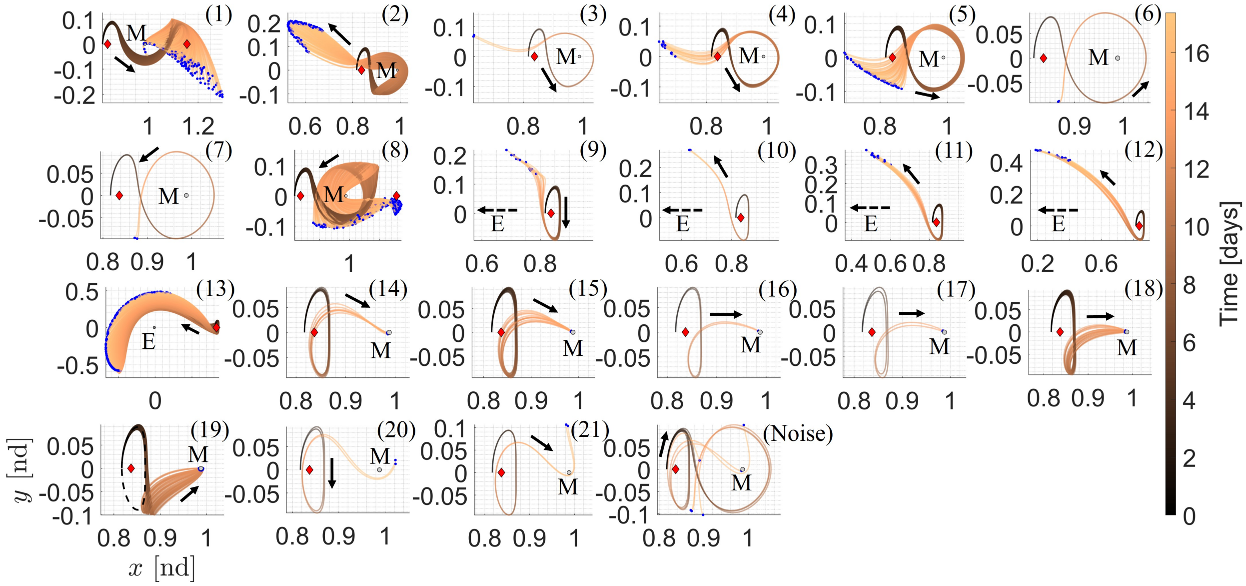

Fig. 9 Final summary of trajectories generated for up to 17.3 days from a local point cloud of initial states along the selected L_1 Lyapunov orbit.

and 21 in the summary but do not lie in sufficiently dense regions to be joined with those groups. Accordingly, the presented data-driven process automatically summarizes the generated set of trajectories via detailed groups from localized regions of high density. This summary can extract key motion types, reduce the complexity of manually analyzing a set of trajectories, or potentially enable automation in support of path-planning or orbit prediction tasks.

G. Summary of Governing Parameters and Decisions

A set of parameters and decisions are required to govern the data-driven categorization of a set of trajectories that begin from a local region of the phase space. Table 1 summarizes these parameters and decisions and lists their influence on the resulting summary. Alternative parameters or decisions can be made throughout the trajectory summarization and clustering process to better fit a desired application or increase the level of differentiation between trajectories.

Although the CR3BP is used to demonstrate the proof of concept in this Note, it is expected that the clustering-based method presented in this Section would be useful for categorizing trajectories generated in higher-fidelity models of a three-body system with little modification. For this proof of concept, the CR3BP offers a sufficiently diverse and complex solution space while supporting fast trajectory generation without reliance on additional parameters, such as an initial epoch or spacecraft properties, required by higher fidelity models. Furthermore, the presented approach does not include assumptions based on the selected system or dynamical model fidelity.

Table 1 Parameters and decisions governing the data-driven categorization process

Parameter/Decision	Step(s)	Influence on categorization process
Step size between initial conditions	0	Influences the number of initial condition samples.
Cloud boundaries	0	Determines the size of the region of initial states centered on the reference state.
Integration time	1	Allows distinct geometries to emerge, if they exist. Can increase the required number of samples for sufficient representation of solution geometry.
Discrete sample selection	2	Defines placement of discrete samples along continuous trajectories. Influences fidelity and characteristics of the solution geometry that are captured by the selected states.
Number of discrete states, p	2	Influences fidelity of discrete representation of trajectories and dimension of feature vector.
Feature selection	2	Influences similarity assessment across all trajectories within the set and characteristics that will be reflected in the resulting clusters.
Clustering algorithm	3, 5	Influences the recovered clusters.
$N_{minCore}$	3, 5	Governs core distance calculation that is used to estimate density across the dataset.
$N_{minClust}$	3, 5	Specifies minimum number of members in a single cluster.
ϵ	3, 5	Defines minimum desired resolution for separating two types of trajectories.

V. Results

The data-driven categorization process described in the previous section is used to summarize sets of trajectories generated from local point clouds of initial conditions near three additional reference states in the Earth-Moon CR3BP. These reference states are selected to produce sets of trajectories with different characteristics or itineraries to support exploring both the generalizability and limitations of the presented data-driven categorization process. For all scenarios, the same HDBSCAN input parameters used for the L_1 Lyapunov orbit scenario in the previous section are used: $N_{minCore} = 4$, $N_{minClust} = 5$, $\alpha = 5$ degrees, and $\epsilon_{thresh} = 10^{-3}\sqrt{p}$ in Step 4; and $N_{minCore} = 1$, $N_{minClust} = 2$, $\epsilon = 0$ in Step 5. Table 2 summarizes the reference states in each scenario.

Table 2 Reference states for each selected scenario

Case	Reference Orbit	C_J	$\boldsymbol{x}_{PO} = [\boldsymbol{r}_{PO}, \boldsymbol{v}_{PO}]^T$ [nd]
1	11.994-day L_1 northern halo	3.1486	$\boldsymbol{r}_{PO}$ = [0.824125682194, 0, 0.0566946270474] $\boldsymbol{v}_{PO}$ = [0, 0.167128773665, 0]
2	7.9615-day L_1 NRHO	2.9947	$\boldsymbol{r}_{PO}$ = [0.988454510548, -0.00114952066778, 0.00705658766736] $\boldsymbol{v}_{PO}$ = [-0.0150237631700, -1.82248741510, -0.148294929894]
3	10.915-day DPO	3.1700	$\boldsymbol{r}_{PO}$ = [1.02454653948, 0.0541769443650, 0] $\boldsymbol{v}_{PO}$ = [-0.380916147888, -0.110095105627, 0]

A. Summarized Motions Near L_1 Halo Orbits

1. L_1 Northern Halo Orbit

For this case, a reference state is selected at the leftmost crossing of the xz-plane along the L_1 northern halo orbit depicted with a dashed black curve in cluster 7 of Fig. 10. The local point cloud of initial states are centered on this reference state. The associated trajectories are generated for 17.3 days and described by $p = 12$ states. The data-driven categorization process identifies 15 clusters composed of 1,677 trajectories as displayed in Fig. 10. This figure uses a similar configuration as Fig. 9. However, in this spatial scenario, each subfigure shows an individual cluster projected onto the $xy-$, $xz-$, and $yz-$planes for clarity. There are seven noise points that do not appear in the summary.

Figure 10 reveals that solutions either impact the Moon, remain near the reference orbit, revolve around the Moon with different geometries, or depart to the Earth vicinity. These key motion types are separated across distinct clusters. In some cases, such as cluster 3, a single motion type captures a continuous evolution of a large set of trajectories over the selected time interval. However, propagating these trajectories for a longer time may result in additional geometric differentiation and, potentially, multiple clusters. Alternatively, the trajectories in clusters 4-7 perform one revolution about the Moon. However, due to the selected small values of $N_{minCore}$ and $N_{minClust}$ and slight variations in shape, time, and region of termination, these trajectories are separated into distinct clusters.

2. L_1 Near Rectilinear Halo Orbit

For the selected L_1 near rectilinear halo orbit (NRHO) in the Earth-Moon CR3BP, the reference state is placed near perilune. Then, trajectories are generated from the local point cloud of initial states near this reference state, propagated for 17.3 days and sampled using $p = 10$ states spaced equally in arclength. Applying the data-driven categorization

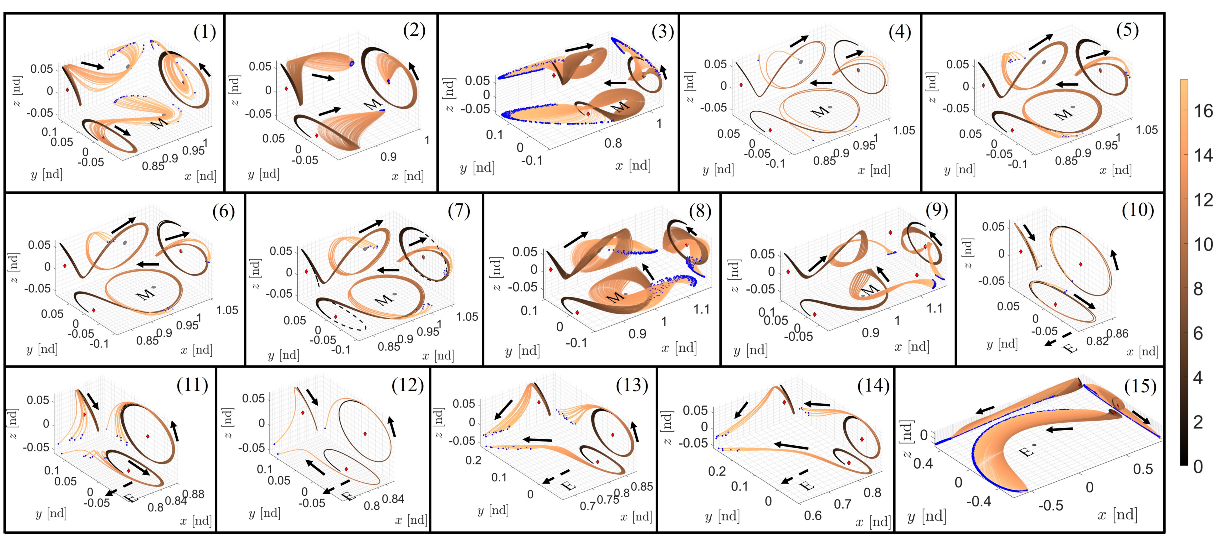

Fig. 10 Fifteen types of 17.3-day trajectories that begin near the selected reference state along a L_1 northern halo orbit at $C_J = 3.1486$ in the Earth-Moon CR3BP.

process to the resulting 3,498 trajectories produces a summary with three distinct motion types and zero ungrouped trajectories. These clusters, along with the selected NRHO, are plotted in Fig. 11 using the same configuration and spatial representation as Fig. 10. In this figure, the second and third clusters separate motions that impact the Moon from those that quickly depart the vicinity of the NRHO with substantially different geometries. The first cluster in Fig. 11 captures trajectories that remain close to the NRHO for the majority of the 17.3 days before finally departing. On either side of the cluster, the trajectories end in different regions of the solution space: either in the Moon vicinity or the Earth vicinity. However, across the cluster, the trajectories continuously evolve between these two behaviors, consistent with the use of a velocity-based feature vector and governing parameters that focus on local assessments of similarity.

B. Summarized Motions Near a Distant Prograde Orbit

Trajectories are propagated from a local point cloud of initial states near a reference state that lies between perilune and apolune along a DPO at a Jacobi constant of 3.1700 in the Earth-Moon CR3BP. In this example, 1,506 trajectories are generated for 26.05 days to allow distinct geometries to emerge and described by $p = 24$ states. All but nine of the trajectories are summarized in the 18 clusters depicted in Fig. 12, with the same representation as Fig. 9.

Across the 18 clusters in Fig. 9, trajectories in each cluster possess a similar geometry and distinct geometries appear in different clusters. For instance, clusters 1-2 capture trajectories that resemble the reference DPO. Clusters 4-9, however, summarize trajectories that depart through the L_2 gateway. These trajectories are likely separated into different clusters due to the evolution of the sampled states as they reach distinct locations in the exterior region. Cluster 3 then captures a transition between these two itineraries. Next, clusters 10-11 capture trajectories that perform a few revolutions around the Moon with two different geometries. Finally, clusters 12-18 contain trajectories with a similar geometry that are distributed across multiple clusters due to localized variations in the 1) path near the first apolune, 2) second perilune passage, and 3) location of the final state. This additional geometric differentiation supplies a more detailed summary of the key motion types.

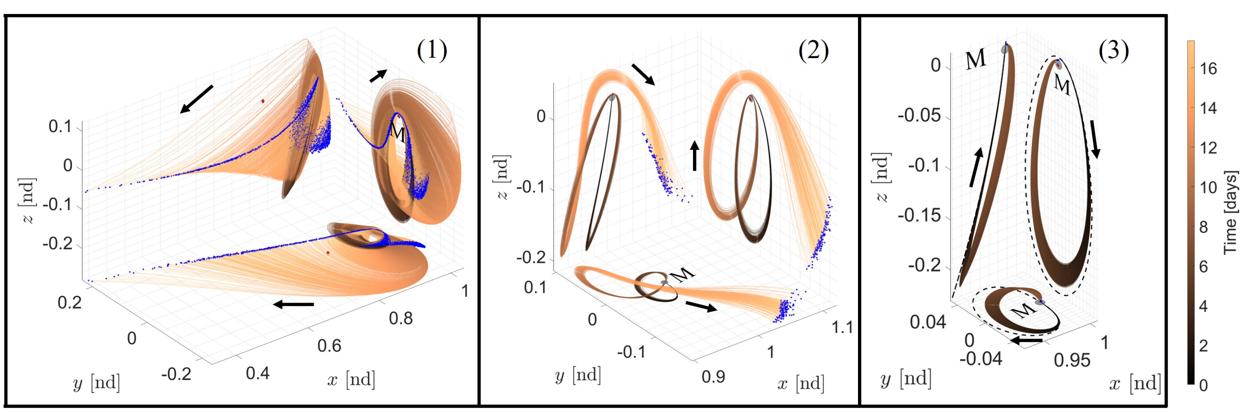

Fig. 11 Three types of 17.3-day trajectories that begin near perilune along the selected L_1 near rectilinear halo orbit at C_J = 2.9947 in the Earth-Moon CR3BP.

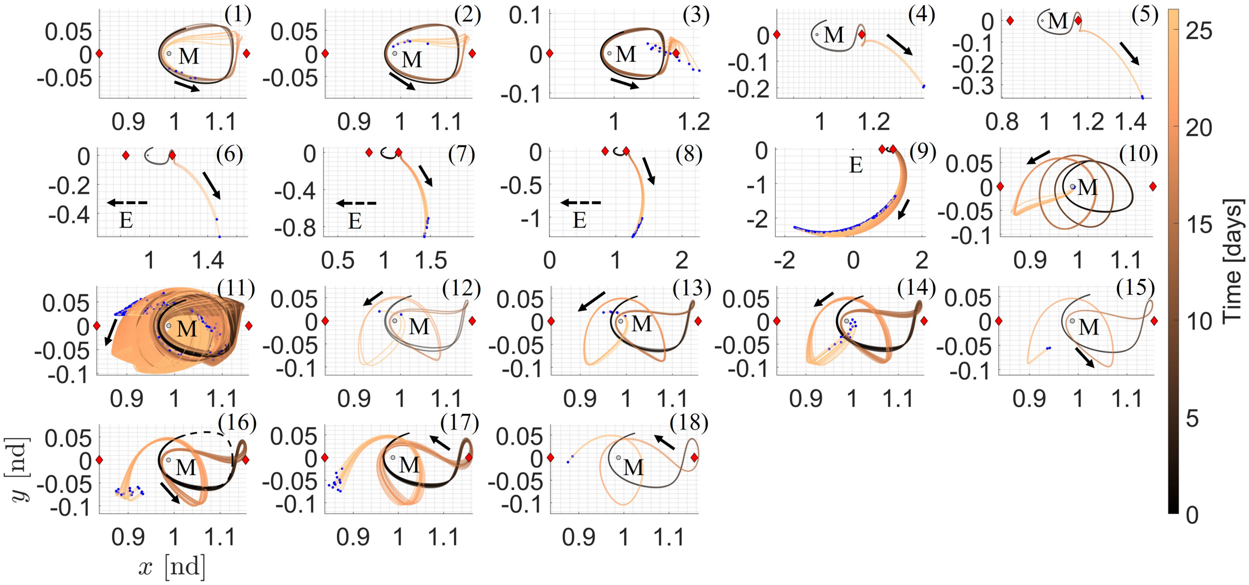

Fig. 12 Distinct types of 26.05-day trajectories beginning near a reference state between perilune and apolune along a DPO at C_J = 3.1700 in the Earth-Moon CR3BP.

VI. Conclusion

In this Note, a clustering-based framework was developed to summarize the types of possible motions generated from a local point cloud of initial conditions within a chaotic multi-body gravitational system. Through application to multiple scenarios in the Earth-Moon CR3BP, this approach was demonstrated to automatically generate groups of trajectories with distinct spatiotemporal evolutions, when these differences exist. In each case, these clusters summarized over 1,000 trajectories via several key motion types. However, some limitations emerged, such as one large cluster including trajectories that continuously evolve across the set to exhibit multiple geometries or one motion type being split into multiple smaller clusters due to variations in the final state impacting sample placement. Alternative feature vectors, sampling schemes, and clustering algorithms may address these limitations in future work. In addition, future application to a wider variety of dynamical models, celestial systems, and initial state sets may support characterizing the quality of the generated summaries and assessing the generalizability of a clustering-based framework. The automatic generation of these summaries may eventually reduce the burden on an analyst or support automation when identifying potential spacecraft motions for space situational awareness or collision-free path-planning.

Funding Sources

This Note was supported in part by a fellowship award under contract FA9550-21-F-0003 through the National Defense Science and Engineering Graduate (NDSEG) Fellowship Program, sponsored by the Air Force Research Laboratory (AFRL), the Office of Naval Research (ONR) and the Army Research Office (ARO). The authors also thank the reviewers for their feedback.

References

[1] National Research Council, *Continuing Kepler's Quest: Assessing Air Force Space Command's Astrodynamics Standards*, The National Academies Press, Washington, D.C., 2012, Chap. 1. https://doi.org/10.17226/13456.

[2] Han, J., Kamber, M., and Pei, J., *Data Mining Concepts and Techniques*, 3rd ed., Morgan Kaufmann Publishers, Waltham, MA, 2012, Chaps. 1, 10. https://doi.org/10.1016/C2009-0-61819-5.

[3] Aggarwal, C. C., and Reddy, C. K., *Data Clustering: Algorithms and Applications*, CRC Press, 2014, Chaps. 1, 3, 15.

[4] Hadjighasem, A., Karrasch, D., Teramoto, H., and Haller, G., "A Spectral Clustering Approach to Lagrangian Vortex Detection," *Physics Review E*, Vol. 93, No. 6, 2016. https://doi.org/10.1103/PhysRevE.93.063107.

[5] Nakhjiri, N., and Villac, B., "Automated Stable Region Generation, Detection, and Representation for Applications to Mission Design," *Journal of Celestial Mechanics and Dynamical Astronomy*, Vol. 123, 2015, pp. 63–83. https://doi.org/10.1007/s10569-015-9629-0.

[6] Haigh, K. Z., Foslien, W., and Guralnik, V., "Data Mining for Space Applications," *Space OPS 2004 Conference*, Montreal, Quebec, Canada, 2004. https://doi.org/10.2514/6.2004-255-106.

[7] Bosanac, N., "Data Mining Approach to Poincaré Maps in Multi-Body Trajectory Design," *Journal of Guidance, Control, and Dynamics*, Vol. 43, No. 6, 2020, pp. 1190–1200. https://doi.org/10.2514/1.G004857.

[8] Bonasera, S., and Bosanac, N., "Applying Data Mining Techniques to Higher-Dimensional Poincaré Maps in the Circular Restricted Three-Body Problem," *Journal of Celestial Mechanics and Dynamical Astronomy*, Vol. 133, No. 51, 2021. https://doi.org/10.1007/s10569-021-10047-3.

[9] Bosanac, N., "Data-Driven Summary of Natural Spacecraft Trajectories in the Earth-Moon System," *2023 AAS/AIAA Astrodynamics Specialist Conference*, Big Sky, MT, 2023.

[10] Smith, T., and Bosanac, N., "Constructing Motion Primitive Sets to Summarize Periodic Orbit Families and Hyperbolic Invariant Manifolds in a Multi-Body System," *Journal of Celestial Mechanics and Dynamical Astronomy*, Vol. 134, No. 7, 2022. https://doi.org/10.1007/s10569-022-10063-x.

[11] Miceli, G. E., Bosanac, N., Mesarch, M. A., Folta, D. C., and Mesarch, R. L., "Clustering Approach to Identifying Low Lunar Frozen Orbits in a High-Fidelity Model," *2023 AAS/AIAA Astrodynamics Specialist Conference*, Big Sky, MT, 2023.

[12] Campello, R. J., Moulavi, D., and Sander, J., "Density-Based Clustering Based on Hierarchical Density Estimates," *Advances in Knowledge Discovery and Data Mining*, Springer, Berlin, Heidelberg, 2013, p. 160–172. https://doi.org/10.1007/978-3-642-37456-2_14.

[13] Campello, R. J. G. B., Moulavi, D., Zimek, A., and Sander, J., "Hierarchical Density Estimates for Data Clustering, Visualization, and Outlier Detection," *ACM Transactions on Knowledge Discovery from Data*, Vol. 10, No. 1, 2015, pp. 1–51. https://doi.org/10.1145/2733381.

[14] Vittaldev, V., Russell, R. P., and Linares, R., "Spacecraft Uncertainty Propagation using Gaussian Mixture Models and Polynomial Chaos Expansions," *Journal of Guidance, Control, and Dynamics*, Vol. 39, No. 9, 2016, pp. 2163–2169. https://doi.org/10.2514/1.G001571.

[15] Sivasankar, P., and Elgohary, T., "Cislunar Orbital Uncertainty Propagation Through the Application of Radial Basis Function Approximation," *AIAA SciTech 2024 Forum*, Orlando, FL, 2024. https://doi.org/10.2514/6.2024-0202.

[16] Bolden, M., Hussein, I., Borowski, H., See, R., and Griggs, E., "Probabilistic Initial Orbit Determination and Object Tracking in Cislunar Space Using Optical Sensors," *Advanced Maui Optical and Space Surveillance Technologies Conference*, Maui, HI, 2022.

[17] Szebehely, V., *Theory of Orbits: The Restricted Problem of Three Bodies*, Academic Press, 1967, Chap. 1.

[18] Koon, W. S., Lo, M. W., Marsden, J. E., and Ross, S. D., *Dynamical Systems, the Three Body Problem and Space Mission Design*, Marsden Books, New York, 2011, Chap. 2.

[19] Doedel, E. J., Paffenroth, R. C., and Keller, H. B., "Computation of Periodic Solutions of Conservative Systems with Application to the 3-Body Problem," *International Journal of Bifurcation and Chaos*, Vol. 13, No. 6, 2003, pp. 1353–1381.

[20] Malzer, C., and Baum, M., "A Hybrid Approach To Hierarchical Density-based Cluster Selection," *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Karlsruhe, Germany, 2020. https://doi.org/10.1109/MFI49285.2020.9235263.

[21] McInnes, L., Healy, J., and Astels, S., "hdbscan: Hierarchical density based clustering," *Journal of Open Source Software*, Vol. 2, No. 11, 2017. https://doi.org/10.21105/joss.00205.

[22] Loyola R, D. G., Pedergnana, M., and Gimeno García, S., "Smart Sampling and Incremental Function Learning for Very Large High Dimensional Data," *Neural Networks*, Vol. 78, 2016, pp. 75–87. https://doi.org/10.1016/j.neunet.2015.09.001.

[23] MathWorks, "MATLAB," https://www.mathworks.com/products/matlab.html, 2025.

[24] Bradley, N., Olikara, Z., Bhaskaran, S., and Young, B., "Cislunar Navigation Accuracy Using Optical Observations of Natural and Artificial Targets," *Journal of Spacecraft and Rockets*, Vol. 57, No. 4, 2020, pp. 777–792. https://doi.org/10.2514/1.A34694.

[25] Fedeler, S., Holzinger, M., and Whitacre, W., "Sensor Tasking in the Cislunar Regime Using Monte Carlo Tree Search," *Advances in Space Research*, Vol. 70, No. 3, 2022, pp. 792–811. https://doi.org/10.1016/j.asr.2022.05.003.

[26] Vallado, D. A., *Fundamentals of Astrodynamics and Applications*, 4^{th} ed., Microcosm Press and Springer, 2013.

[27] Shifrin, T., *Differential Geometry: A First Course in Curves and Surfaces*, University of Georgia, 2023, Chap. 1.

[28] Bosanac, N., "Data-Driven Summary of Motion in an Ephemeris Model of Cislunar Space," *2025 AAS/AIAA Space Flight Mechanics Meeting*, Kaua'i, HI, 2025.

[29] Aurenhammer, F., "Voronoi Diagrams - A Survey of a Fundamental Geometric Data Structure," *ACM Computing Surveys*, Vol. 23, No. 3, 1991, pp. 345–405. https://doi.org/10.1145/116873.116880.