



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Memory Clustering Using Persistent Homology for Multimodality- and Discontinuity-Sensitive Learning of Optimal Control Warm-Starts

**Citation for published version:**

Merkt, WX, Ivan, V, Dinev, T, Havoutis, I & Vijayakumar, S 2021, 'Memory Clustering Using Persistent Homology for Multimodality- and Discontinuity-Sensitive Learning of Optimal Control Warm-Starts', *IEEE Transactions on Robotics*, pp. 1-12. <https://doi.org/10.1109/TRO.2021.3069132>

**Digital Object Identifier (DOI):**

[10.1109/TRO.2021.3069132](https://doi.org/10.1109/TRO.2021.3069132)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

IEEE Transactions on Robotics

**General rights**



Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Memory Clustering using Persistent Homology for Multimodality- and Discontinuity-Sensitive Learning of Optimal Control Warm-starts

Wolfgang Merkt\*  Vladimir Ivan\*  Traiko Dinev  Ioannis Havoutis  Sethu Vijayakumar 

**Abstract**—Shooting methods are an efficient approach to solving nonlinear optimal control problems. As they use local optimization, they exhibit favorable convergence when initialized with a good warm-start but may not converge at all if provided with a poor initial guess. Recent work has focused on providing an initial guess from a learned model trained on samples generated during an offline exploration of the problem space. However, in practice the solutions contain discontinuities introduced by system dynamics or the environment. Additionally, in many cases multiple equally suitable, i.e., multi-modal, solutions exist to solve a problem. Classic learning approaches smooth across the boundary of these discontinuities and thus generalize poorly. In this work, we apply tools from algebraic topology to extract information on the underlying structure of the solution space. In particular, we introduce a method based on persistent homology to automatically cluster the dataset of precomputed solutions to obtain different candidate initial guesses. We then train a Mixture-of-Experts within each cluster to predict state and control trajectories to warm-start the optimal control solver and provide a comparison with modality-agnostic learning. We demonstrate our method on a cart-pole toy problem and a quadrotor avoiding obstacles, and show that clustering samples based on inherent structure improves the warm-start quality.

## I. INTRODUCTION

**O**PTIMAL control can be used to generate highly dynamic motions and behaviors by specifying objectives composed of desirable characteristics with dynamics taken as constraints. In particular, it allows composing complex maneuvers by working over a long time horizon. In contrast, instantaneous control methods such as inverse dynamics are unable to solve such problems. Examples of this include swing-up of an under-actuated cartpole or jumps and front-flips on legged platforms [1]. However, optimal control methods can take longer to converge if not warm-started and can easily get stuck in a local minimum if provided with a poor initial guess. Previous work has focused on exploration of the parameterized optimal

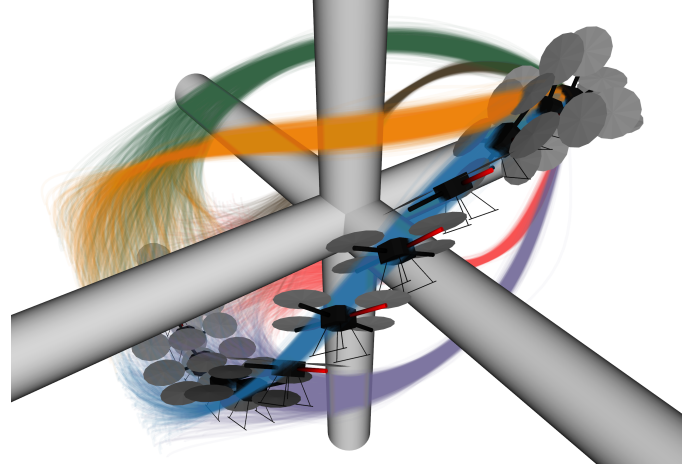


Figure 1. Illustration of discontinuity and multi-modality: Quadcopter flying in a complex environment: The different classes of trajectories from start to goal cannot be continuously deformed into each one another. This violates the continuous map assumption between problem parametrization and solution output  $f : X \rightarrow Y$  which is core to function approximation. The number of clusters and class memberships are automatically extracted using our method.

control problem in an offline computation process to gather experience to initialize from during runtime [2]–[10].

When considering optimal control problems, one can observe two cases that make it challenging to directly use learning as a way of compressing and generalizing across optimal solution samples: First, *discontinuity*, where similar problems (in parameter space) yield vastly different optimal solutions. Examples of this can be seen in a) the phase space plot of optimal solutions for a pendulum or cartpole swing-up task and b) with discontinuities in solution paths introduced by environment obstacles. Second, *multi-modality*, where multiple equally optimal solutions to a problem exist. A prominent example is the ability to traverse around an obstacle in multiple ways as seen in Figure 1.<sup>1</sup>

Both discontinuity and multi-modality can greatly impact the quality of prediction obtained using function approximation as regressors smooth across the boundaries between clusters or modalities. This is expected as efficient function approximation

<sup>1</sup>We note that in some cases one mode may be less optimal than another, however, as most work relies on locally optimal samples generated using a stochastic precomputation process, this may not be known a priori. Additionally, storing multiple modes can increase warm-start success and robustness as one mode may, for instance, be obstructed by an obstacle.

\* The first two authors contributed equally to this work.

This research was supported by (1) the European Commission under the Horizon 2020 project Memory of Motion (MEMMO, project ID: 780684), (2) the Engineering and Physical Sciences Research Council (EPSRC) UK RAI Hub for Offshore Robotics for Certification of Assets (ORCA, grant reference EP/R026173/1), and (3) the EPSRC Centre for Doctoral Training in Robotics and Autonomous Systems (EPSRC, grant reference EP/L016834/1). (Corresponding author: Wolfgang Merkt.)

Wolfgang Merkt and Ioannis Havoutis are with the Oxford Robotics Institute, University of Oxford, Oxford, OX2 6NN, U.K. (e-mail: wolfgang@robots.ox.ac.uk; ioannis@robots.ox.ac.uk).

Vladimir Ivan, Traiko Dinev, and Sethu Vijayakumar are with the School of Informatics, The University of Edinburgh, Edinburgh EH8 9AB, U.K. (e-mail: v.ivan@ed.ac.uk; traiko.dinev@ed.ac.uk; sethu.vijayakumar@ed.ac.uk).

methods such as Locally Weighted Projection Regression (LWPR) [11] and Gaussian Process Regression (GPR) [12] assume uni-modal distributions and continuity. One workaround to avoid multi-modality is to bias the sampling/exploration stage to include/enforce only one modality (e.g., [7]). Alternatively, we can consider machine learning models that can handle multi-modality and discontinuity directly [13]. These are often a combination of multiple local models that each model one continuous cluster of data well. Clustering time-series data and learning trajectories has been widely explored under the paradigm of Programming by Demonstration (PbD) to extract a set of alternative, feasible solutions from a dataset of demonstration trajectories. For input partitioning, hierarchical clustering has previously been applied using simple distance-based geometric approaches [14] and Hidden Markov Models (HMMs) [15], [16]. The former required heuristics to be set while the latter was reported by [14] to be unstable. Previous work circumventing this issue used trajectory libraries queried with nearest neighbor [2]–[4], [8], [17], a hyper-local model that returns the closest neighbor without any interpolation. Thus, the returned solution is valid (i.e., does not violate any constraints for a similar problem), but does not generalize between samples (and also does not compress the original dataset). In order to improve success rates for handling discontinuous and multi-modal distributions directly, Mixture-of-Experts (MoE) [18] and Product-of-Experts (PoE) [19] systems have been proposed. In MoE, a gating function or network determines which expert will provide the best output. The rationale is that a regressor trained exclusively on a continuous subset of the data will do well when queried within that subset but poorly outside. Traditionally, MoE systems first partition the data based on the similarity of the input-output mapping and then train experts individually on subsets of data. If a clustering is not known, joint training using a loss function that encourages both specialization and cooperation between the local experts, and thus automatically assigns samples to classes, can be employed [18]. PoE methods, on the other hand, combine the output of multiple probabilistic models to form a prediction. An example is the Gaussian Mixture Models (GMMs) which can capture multi-modality directly, however, they require information on the number of classes present or the tuning of hyper-parameters. [20], for instance, use a Dirichlet process which can possibly represent an infinite number of clusters, while [21] applies a Dirichlet distribution prior with a fixed number of clusters.

The authors of [9] used a MoE approach for discontinuity-sensitive learning of initialization seeds. They applied k-means clustering informed by expert knowledge on the input-output relationship such as periodicity of angles or Lagrange multipliers of constraints. They then learned a MoE system and applied it on a pendulum, 2D car, and a quadrotor with a single spherical obstacle.

We argue that while these approaches may work well on small problems and with datasets where a system designer’s intuition is readily available, it is challenging to extract heuristics for labeling data (or to extract the number of clusters) on higher dimensional tasks. Furthermore, due to the stochastic nature of the exploration stage, the dataset may not include samples of all topologically distinct classes or modalities.

The authors of [22] took a different approach to cluster trajectories. They used filtrations of simplicial complexes and persistent homology for modeling trajectories in configuration spaces. They then used the persistency of simplicial complexes to classify trajectories with fixed start and end points. This method looks at the changes in topology across different scales and identifies at which scale topological features (connected components and holes) appear and disappear. The theory behind this approach has been studied in [23], and computationally efficient algorithms have been proposed in [24]. These were further improved for computational speed and memory efficiency in [25] and [26]. Motivated by this, we aim to automatically extract information on the underlying solution space.

Similar to [22], we use filtrations of simplicial complexes to extract information about the topological structure of trajectories in a dataset—in particular, to reason about the classification of trajectories using the first homology group. In contrast to [22], we use Vietoris-Rips complexes for scalability and introduce filtrations based on trajectory segment distances which allows us to scale to larger datasets and further offers adaptability to particular applications through the selection of the segment distance metric. In our experiments, we show scalability to 78-dimensional trajectories and analyze the scalability of our method. Our method focuses on extracting equivalence classes based on the persistent topological structure of the sample data to train a memory-of-motion to initialize/warm-start optimal control planning. We evaluate our method on initializing Differential Dynamic Programming (DDP)-style optimal control methods for cartpole and quadrotor navigation tasks. We further explore scalability through a humanoid manipulation example. The key benefit here is to allow incorporating global information with local trajectory optimization methods [22].

In this paper, we make the following contributions:

- 1) We introduce a novel method for clustering time-series data based on persistent homology to identify multiple modes and discontinuities among continuous trajectories within a precomputed dataset.
- 2) We improve scalability to large datasets of time-series data by leveraging a segment-to-segment distance for filtration and use a pairwise trajectory distance for clustering.
- 3) We propose a procedure based on half-life to automatically extract the number of clusters from the persistence of cohomology groups.
- 4) We demonstrate that this method can be applied on state spaces of different dimensions and is also applicable to high-dimensional task spaces.
- 5) We evaluate the scalability of the method w.r.t. the number of trajectory samples and the number of time steps per sample.
- 6) We show that our method outperforms methods that do not exploit the clustering for warm-starting and we demonstrate this on a dynamic optimal control task in a complex environment.

To the best of our knowledge, we are the first to describe a fully automatic trajectory clustering pipeline based on persistent homology; and the use of these tools on a dataset of motions

of highly dynamic systems such as a cartpole and a quadrotor as well as to warm-start optimization solvers.

## II. OPTIMAL CONTROL

We focus on discrete-time, finite-horizon nonlinear optimal control. Consider a dynamical system for which we aim to find a policy  $\mathbf{u} = \pi(\mathbf{x})$  that minimizes a cost function

$$J = \ell_f(\mathbf{x}_T) + \sum_{t=1}^{T-1} \ell(\mathbf{x}_t, \mathbf{u}_t). \quad (1)$$

Starting from an initial state  $\mathbf{x}_0$ , the system evolves according to the state transition function  $\mathbf{x}_{t+1} = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t)$ , which incorporates the differential system dynamics and an integration scheme.  $\ell_f(\mathbf{x}_T)$  denotes the state cost at the end of the horizon and  $\ell(\mathbf{x}_t, \mathbf{u}_t)$  the general running cost. We discretize in  $T$  time steps and minimize  $J$  to obtain a sequence of controls  $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{T-1}]$ , where all controls are bounded with upper and lower limits  $\mathbf{u}_t \in [\underline{\mathbf{u}}, \bar{\mathbf{u}}]$ . We refer to the minimal cost for a state  $\mathbf{x}$  at time  $t$  as the cost-to-go  $V(\mathbf{x}_t, T)$ .

For the purpose of this work, we will solve the above optimal control problem using a direct-indirect hybridization approach [27], a recent extension to shooting methods with improved globalization (i.e., robustness to poor initialization) and the ability to provide initial guesses for both state and control trajectories. Because of this, we only consider general costs and bounded control constraints. We do not explicitly use constraints on the states but include these as cost terms or enforce them in the forward simulation.

### A. Differential Dynamic Programming

DDP is a second-order shooting method optimizing only over the unconstrained control space displaying quadratic convergence [28], [29]. As a gradient descent method, it uses locally-quadratic approximations of the dynamics and cost functions. DDP alternates between a backward pass on a reference trajectory to generate a new sequence of local feedback control laws, and a forward pass computing the new state trajectory. Following [30], if  $Q$  is the variation of the cost-to-go  $V$  (with the subsequent cost-to-go denoted as  $V'$ ), we can expand its variation to second-order using a Taylor series as:

$$Q(\delta\mathbf{x}, \delta\mathbf{u}) \approx \frac{1}{2} \begin{bmatrix} 1 \\ \delta\mathbf{x} \\ \delta\mathbf{u} \end{bmatrix}^T \begin{bmatrix} 0 & Q_x^T & Q_u^T \\ Q_x & Q_{xx} & Q_{xu} \\ Q_u & Q_{ux} & Q_{uu} \end{bmatrix} \begin{bmatrix} 1 \\ \delta\mathbf{x} \\ \delta\mathbf{u} \end{bmatrix}, \quad (2)$$

where the individual terms are:

$$Q_x = \ell_x + \mathbf{f}_x^T V'_x \quad (3)$$

$$Q_u = \ell_u + \mathbf{f}_u^T V'_x \quad (4)$$

$$Q_{xx} = \ell_{xx} + \mathbf{f}_x^T V'_{xx} \mathbf{f}_x + V'_x \cdot \mathbf{f}_{xx} \quad (5)$$

$$Q_{uu} = \ell_{uu} + \mathbf{f}_u^T V'_{xx} \mathbf{f}_u + V'_x \cdot \mathbf{f}_{uu} \quad (6)$$

$$Q_{ux} = \ell_{ux} + \mathbf{f}_u^T V'_{xx} \mathbf{f}_x + V'_x \cdot \mathbf{f}_{ux}. \quad (7)$$

Solving for the optimal change in control  $\delta\mathbf{u}^*$  given a change in state  $\delta\mathbf{x}$  we have

$$\delta\mathbf{u}^* = \arg \min_{\delta\mathbf{u}} Q(\delta\mathbf{x}, \delta\mathbf{u}) \quad (8)$$

$$= -Q_{uu}^{-1}(Q_u + Q_{ux}\delta\mathbf{x}), \quad (9)$$

$$= \mathbf{k} + \mathbf{K}\delta\mathbf{x} \quad (10)$$

where  $\mathbf{k}$  is the feed-forward term and  $\mathbf{K}$  the feedback gain matrix. Using this result, Equation (2) can be solved for the quadratic model of the value change:

$$\Delta V = -\frac{1}{2} Q_u Q_{uu}^{-1} Q_u \quad (11)$$

$$V_x = Q_x - Q_u Q_{uu}^{-1} Q_{ux} \quad (12)$$

$$V_{xx} = Q_{xx} - Q_{xu} Q_{uu}^{-1} Q_{ux}. \quad (13)$$

We apply regularization on the states and controls following [30, Ch. 2] to ensure that the problem remains numerically well-conditioned. To directly incorporate bound constraints on the controls without sacrificing convergence, [31] introduced the use of a projected-Newton class active-set Quadratic Programming (QP) method. Line search with variable step sizes is used to achieve convergence. Recent improvements to DDP introduced a hybridization called Feasibility-driven Differential Dynamic Programming (FDDP) which is comparable to multiple shooting and which has shown greater globalization from poor initial guesses [1]. Here, we use a recent variant that combines [1] and [31], Control-limited Feasibility-driven Differential Dynamic Programming (BoxFDDP), to explicitly handle control bounds [27].

The resulting control sequence and trajectory will be locally optimal. We obtain a set of locally optimal solutions by initializing DDP with different initial control sequences and from different initial states.

Having obtained a set of solutions to parameterized optimal control problems as a dataset, we are interested in learning a mapping  $f : X \rightarrow Y$  from the problem encoding/parameterization to control trajectory output. There is, however, no guarantee that the dataset is uni-modal nor that any of the solutions is the global optimum. Therefore, to train a warm-start model that can initialize the solver close to a global optimum, we are first interested in grouping the trajectories into uni-modal clusters.

## III. CLUSTERING USING PERSISTENT HOMOLOGY

Our objective is for all trajectories within a cluster to not have any topological holes (e.g., caused by obstacles) between them that would cause the warm-start to get stuck in a local minimum. In topology, such property can be described using homotopy. A homotopy between two trajectories exists if we can define a smooth continuous function that transforms one trajectory into another [32]. Defining homotopies is an active subject of research and it is considered a difficult problem in the field of topology. However, topological features such as clusters and holes define homology equivalences that are commonly used to approximate homotopy equivalences. Here, efficient algorithms exist to compute homology groups of simplicial complexes (see below). Knowledge of these homology groups

allows us to reason about the global properties of a space based on local computations [33].

To analyze a dataset, we consider it as a set of points associated with a distance metric. The points are the state and control trajectories as computed by an optimal control method, discretized in time. From these points, we are particularly interested in computing the invariant *cohomology* of the output space. This can be interpreted as the number of “holes” that would separate the data into clusters. A key building block of algebraic topology are *simplicial complexes* made up of simplices: Points are 0-simplices, edges/lines between two points form a 1-simplex, three points forming a triangular face are a 2-simplex, and so on for higher-dimensional simplices. During filtration, we increase the distance (or scale) parameter  $r$ . For each point in the dataset, we connect all points within distance  $r$  to form simplices. We record when simplices emerge, e.g. two points (0-simplices) get connected by an edge, forming a 1-simplex. This event is also called the *birth* of a simplex. At the same time, one of the 0-simplices disappears as it merges into the larger connected component (the edge). This is marked as a *death* of the shorter-lived simplex. We plot these events on the persistence diagram by a dot along the *Birth* and *Death* axis. The events are color-coded by the dimension of the simplex or complex that died in the event (0-simplices -  $H_0$ , 1-simplices -  $H_1$ , etc.) As the value of  $r$  increases, we track how simplices get merged into larger complexes and the persistence of the data is revealed by the distance of the death event from the birth (along the vertical axis). Alternatively, we can plot a persistence diagram as in Figure 2c where the  $x$ -axis denotes the *Birth* and the  $y$ -axis shows the *Lifetime* (Lifetime = Death - Birth). Invariant features of the underlying dataset have a long persistence/lifetime (are higher up) and can thus be read off the homology diagram. Alternatively, these features can be visualized using a barcode diagram [34]. Here, the rank of the zeroth-dimensional homology group ( $H_0$ ) corresponds to the number of connected components, while the rank of the first-dimensional homology group ( $H_1$ , the number of one-dimensional “holes”) allows us to reason about the number of clusters in the data.

### A. Dealing with time-series data

As the trajectories represent time-series data (the state or controls at time steps  $t_1, \dots, t_T$ ), we cannot directly use them as a dataset for filtration.

Suppose we have  $N$  samples of  $M$ -dimensional time-series data with  $T$  time steps. As we are interested in the structure of the  $M$ -dimensional space, a naïve point-cloud approach would be to stack the samples into an  $NT$ -dimensional vector to create a pairwise distance matrix of dimension  $(NT) \times (NT)$ . However, we have to choose a large number of samples and time steps to cover the solution space sufficiently densely. One way to reduce the size of the distance matrix without sacrificing the coverage is to compute the distance between trajectory segments instead of the trajectory knot points. We propose to use linear segments and calculate the distance as proposed in [35]. This allows us to approximate the distance matrix of a set of trajectories with relatively few time steps which decreases

the size of the distance matrix to  $(N(T - 1)) \times (N(T - 1))$ . While this change does not immediately reduce the size of the distance matrix significantly, it allows us to adaptively replace several data points that are approximately linear with a single line segment (i.e., adaptive numbers of  $T$ ).

After we obtain a full distance matrix, we post-process it by incorporating explicit connectivity information from the trajectories: We explicitly set the distance for subsequent time steps to zero as suggested by [22]. We can do the same for connected start and end states. This further ensures that all samples are connected into one  $H_0$  group and works similarly to the common start and end points proposed by [22]. As a result, we always get a single  $H_0$  component in all our persistence plots. We then apply filtration to the post-processed distance matrix to extract the persistent homology groups.

To illustrate, we apply this process to a toy problem, see Figure 2. We generated two pairs of trajectories using sine waves. Each trajectory is a sequence of 2D positions and corresponding 2D velocities, resulting in a trajectory in  $\mathbb{R}^4$ . Figure 2a shows the position trajectory. The first pair of trajectories are touching in the middle, where both the position and velocity equal to zero. The second pair is crossing in the middle with the position being zero but the velocities being large. We have designed the trajectory such that the velocity along the  $x$ -axis is constant. This allows us to visualize the 4D phase-space plot by dropping the velocity along the  $x$ -axis and plotting the velocity along the  $y$ -axis as the  $z$ -coordinate in the 3D plot in Figure 2b. This plot shows how the trajectories that are touching create two loops, while the trajectories that are crossing create only one loop in the phase-space. The persistent homology plot in Figure 2c shows this topology. On the left, we see two  $H_1$  groups (orange dots), and on the right, one  $H_1$  group that corresponds to the number of loops in the phase-space of each pair of trajectories. There are also several short-lived  $H_1$  groups close to the bottom of the graph. These are small holes that correspond to the aliasing artifacts. They depend on the resolution at which we sampled the sine waves and can be considered as noise. In both cases, there is one  $H_0$  group (blue dot) that captures the single connected component created by connecting the ends of the two trajectories.

We are interested in  $H_1$  groups with early birth (furthest to the left on the diagram) and long lifespan (higher up). These translate into persistent features in the data that we want to preserve. We can additionally extract the separating distance as  $> 0.3$  (above where the short-lived groups disappear but below the long-lived groups appear).

Note, however, that the proposed process considers distance matrices in their dense form, which can very quickly exhaust memory during computation.<sup>2</sup> In practice, we frequently reduce the sampling frequency along the time dimension to make the problems tractable. The toy example trajectory can be sampled at as few as  $T = 5$  time steps without changing the topology of the dataset.

An alternative approach would be to use dimensionality reduction tools or alternate representation with embeddings

<sup>2</sup>We have explored the use of filtration using sparse distance matrices [26], however, the temporal structure of time-series data is not compatible with this technique.

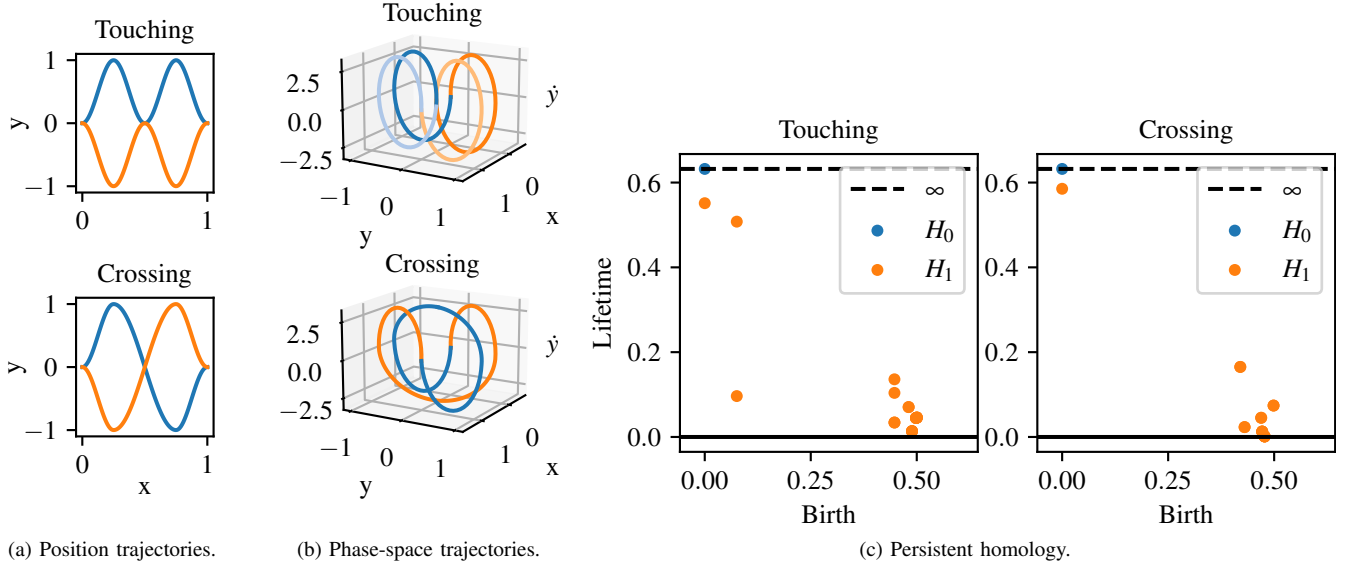


Figure 2. Toy example comparing topology of a pair of trajectories (blue and orange) that touch each other at zero velocity (top) with a pair that crosses (bottom). Figure 2b shows the phase-space plot where the trajectories show a double loop and a single loop respectively. These loops are detected as two and one  $H_1$  groups in the top left corner of the persistent homology plot in Figure 2c.

before filtration. Appropriate alternate representations are crucial as they can significantly change the topology of the space. This is especially crucial when dealing with time-series data that includes state derivatives and dimensions with various units. This is common with robot joint position and velocities having different scales, and with linear joints, angular joints, and the floating base being measured in different units. We manually set the relative scaling and use the persistent homology to analyze the topology of the resulting space. In the toy problem, this corresponds to scaling the  $z$ -axis in Figure 2b. Once we have selected the scaling, we proceed to use the same tools for clustering the dataset. We explore this concept further using alternate space representations in Section V-D.

### B. Extraction of number of clusters from persistence of cohomology groups

The lifetime (persistence) of simplicial complexes is commonly visualized in a persistence diagram as in Figure 2c. To automatically extract the number of significant separating radii, we propose to use a heuristic based on half-life between subsequent lifetimes in an ordered list. We show this procedure in Algorithm 1. Two parameters are the cut-off ratio between subsequent lifetimes (we use 0.8) and the minimum lifetime (distance) to filter out short-lived simplicial complexes (we use 0.1).

We ignore the  $H_0$  group since all trajectories always belong to a single connected component due to how we treat the time-series data. We count the number of  $H_1$  groups and compute the number of clusters as  $H_1 + 1$ . We use this number as an input to the clustering algorithm.

### C. Clustering

Knowing the number of classes in the dataset, we now need to apply labels to each trajectory sample. To achieve

**Algorithm 1** Algorithm for extracting the number of clusters based on the first homology group.

**Input:**  $H_1$  groups; cutoff\_ratio; min\_lifetime

**Output:** num\_classes

```

1: num_classes = 1
2: lifetimes = SortDescending(Deaths - Births)
3: previous_lifetime = infinity
4: for lifetime  $\in$  lifetimes do
5:   if (lifetime < min_lifetime or
        lifetime < cutoff_ratio * previous_lifetime) then
6:     return num_classes
7:   else
8:     num_classes+ = 1
9:     previous_lifetime = lifetime
10:  end if
11: end for
12: return num_classes

```

this, we create a trajectory-wise distance matrix of size  $D_{\text{pairwise-trajectories}} \in \mathbb{R}^{N \times N}$ . To create it, we apply the process using the trajectory segment distance and post-processing for time-series data described in Section III-A for each pair of trajectories  $i, j$  to obtain a pairwise post-processed distance matrix  $D_{ij}$ . This is the same distance matrix we use for computing the persistent homology. We then extract the maximum of the minimum distances across trajectory segments  $D_{ij}$  and store it as the representative distance between trajectories  $i$  and  $j$  in  $D_{\text{pairwise-trajectories}}$ . This metric has a topological meaning. Since we already enforce the distance between subsequent time steps and the start and end segments to be zero, the lifetime of a  $H_1$  group is proportional to the furthest distance between two trajectories computed using our proxy metric. The trajectory-wise distance matrix, therefore, serves as an approximation of a metric based on the persistent

homology of a pair of trajectories. As a result, it encodes how close the two trajectories are to being in the same homotopy class.

Finally, using the number of clusters from Algorithm 1 and with the precomputed distance matrix  $D_{\text{pairwise-trajectories}}$ , we use agglomerative clustering with the single linkage method to assign class labels to the original dataset.

#### IV. MIXTURE OF EXPERTS

Whether, and how quickly, the optimization solver converges from a given initial condition and parameterized goal setting depends significantly on the quality of the initial guess. If we can provide a good initial guess to the optimization solver, considerable speed-up for the convergence can be achieved [8]. Having separated the data into clusters with continuous input-output relationships, we now propose a Mixture-of-Experts approach to predict the state and control trajectories given new initial conditions.<sup>3</sup> We propose to use  $k$  experts, where  $k$  is the number of clusters identified in Section III-B, and a gating network to decide which expert to query. Note, that multiple experts can also be explored independently and simultaneously, similar to ensemble methods. Here, we use simple Multi-Layer Perceptron (MLP) models with ReLu activation for each expert and train using Adam [36]. We did not tune the hyper-parameters apart from ensuring that the cumulative capacity of each of the compared methods matches.

#### V. EVALUATION

We test our methodology on optimal control tasks using the cartpole and quadrotor dynamic models. We implement the optimal control problem, system dynamics, and BoxFDDP solver in the Extensible Optimization Toolset (EXOTica) [37]. For topological analysis, we compute the persistence cohomology of the dataset using the efficient Ripser library [38], [39]. All evaluations were carried out on a laptop using a single core of an Intel Core i7-9850H CPU at 4.2 GHz and 64 GB 2933 MHz memory.

We provide our implementation, datasets, and notebooks as open-source for reproducing our results.<sup>4</sup>

##### A. Scalability

First, we look at the scalability of our proposed method to identify the number of clusters (Section III-A) with i) the number of samples in the dataset ( $N$ ), ii) the length of the trajectories ( $T$ ), and iii) the dimension of the state space. The size of the distance matrix used for filtration scales with  $N$  and  $T$ . Hence, we expect the computational complexity to scale with  $\mathcal{O}((N(T-1))^2)$ . We use the cartpole dataset from Section V-B and cubic interpolation for time re-sampling to evaluate scalability with  $N$  and  $T$ . The results are presented in Figure 3 and show filtration time scaling as expected. Note, the use of line segment distances enables us to achieve scalability

<sup>3</sup>Traditionally, shooting methods can only be initialized using control trajectories. Using FDDP-derived solvers, we can provide both state and control trajectories as an initial guess.

<sup>4</sup>[https://github.com/wxmerk/topological\\_memory\\_clustering](https://github.com/wxmerk/topological_memory_clustering)

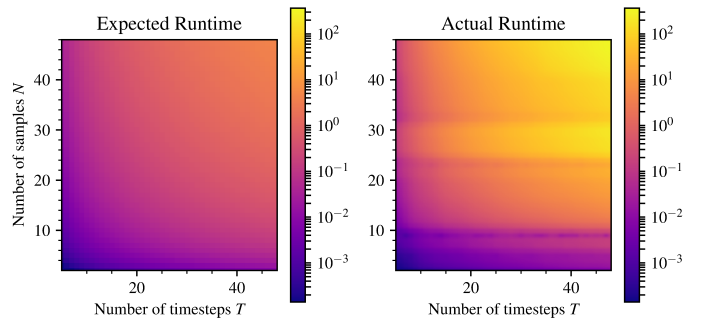
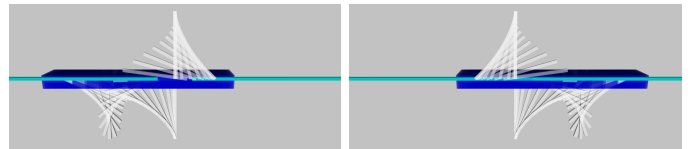


Figure 3. *Left*: Expected computation time (s) scaling with  $\mathcal{O}((N(T-1))^2)$ . *Right*: Computation times from filtration of datasets with varying  $N$  and  $T$ . We use the cartpole dataset from Section V-B and cubic interpolation for time re-sampling.

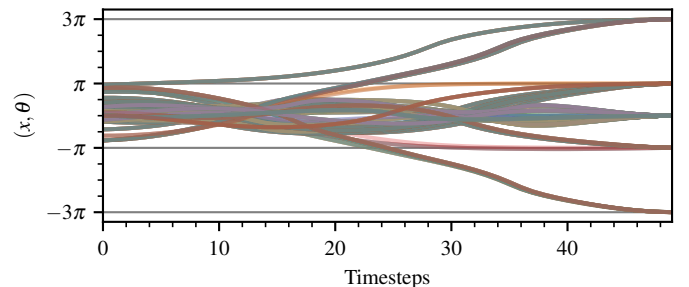
to larger datasets by replacing densely sampled trajectory sub-segments with a single segment-wise distance (e.g., a piecewise linear approximation).

##### B. Cartpole swing-up

The cartpole is a dynamic system where a pole is mounted with an unactuated hinge joint on a cart that travels on a rail. It uses horizontal forces as controls  $\mathbf{u}$ . Due to control limits and as an under-actuated system, the cartpole is a canonical task for nonlinear optimal control as the cart needs to gather energy in order to be able to swing up. We model the system following [40] with the slider position denoted as  $x$ , the angle of the pole as  $\theta$ , and the state as  $[x, \theta, \dot{x}, \dot{\theta}]^T$ . We limit the control input to  $\mathbf{u} \in [-10, 10]$  N. The aim is to swing up the pole to the upright position with the cart at the origin and zero final velocity (i.e.,  $\mathbf{x}_{goal} = [0, \pm\pi, 0, 0]^T$ ). Note, that we do not require the pole to reach the final value of  $\pi$ , but any configuration that is upright: It is irrelevant from which side the pole swings up—or whether it completes more than

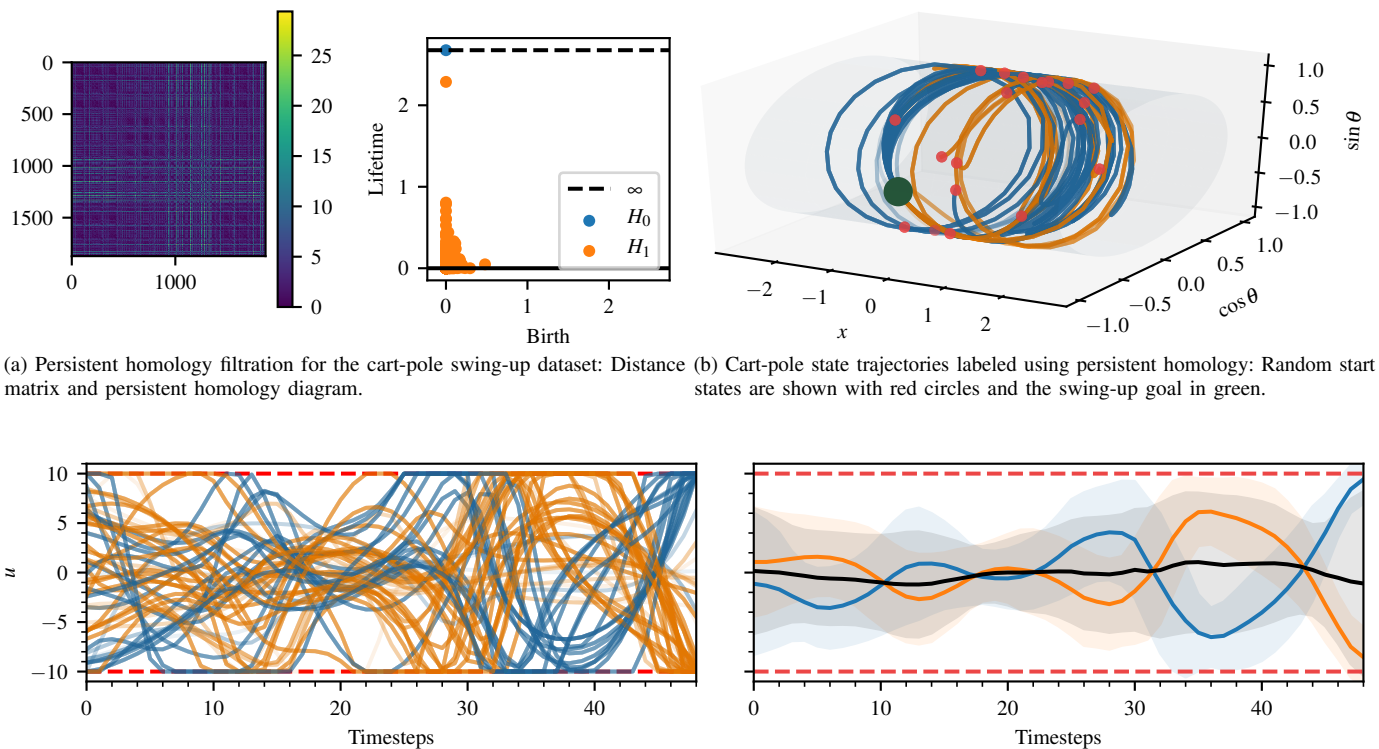


(a) Visualization of the two modes for the swing-up task.



(b) State trajectories for swing-up policies from random initial states.

Figure 4. The cartpole swing-up task: We illustrate the two modes (Figure 4a) and show 500 policies for a cartpole swing-up task from random initial states (Figure 4b). Note, depending on the initial velocity the pole may complete several rotations before stabilizing.



(a) Persistent homology filtration for the cart-pole swing-up dataset: Distance matrix and persistent homology diagram. (b) Cart-pole state trajectories labeled using persistent homology: Random start states are shown with red circles and the swing-up goal in green.

(c) Control trajectories for random start states with class labels assigned from the output of the persistent homology filtration. (d) Mean and standard deviation of the control trajectories for each of the classes (orange, blue) and the dataset (black).

Figure 5. Results on the cart-pole swing-up task: (a) Persistent homology diagram computed from the pre-processed distance matrix showing one hole (two classes) as indicated by a single  $H_1$  group (orange dot in the left top corner). (b) State trajectories labeled using persistent homology. (c) Control trajectories labeled using persistent homology. (d) Mean and standard deviation of control trajectories conditioned by class highlighting that without class-sensitivity, the information cancels out.

one full rotation prior to coming to rest. Therefore, we model the continuous hinge joint using the special orthogonal group  $\mathbb{S}\mathbb{O}(2)$  and represent the state as  $[x, \cos \theta, \sin \theta, \dot{x}, \dot{\theta}]^T$ .

We now consider a swing-up task on the cartpole. From intuition, we can postulate that there are two modes: *swinging up from the left side* and *swinging up from the right side*. Indeed, these two solutions can be seen in Figure 4a.

To build the dataset, we randomly sample start states in the range  $\pm(1 \text{ m}, \pi \text{ rad}, 1 \text{ m s}^{-1}, 0.5\pi \text{ rad s}^{-1})$  and control policies uniformly from  $[-1, 1] \text{ N}$ . We solve the optimal control problem using BoxFDDP until we have  $N = 10$  solutions per random start state. The state trajectories in Figure 4b show that the pole may complete several full rotations prior to reaching a zero velocity and getting into the upright pose.

We now aim to extract (a) the number of topological classes contained in the dataset and (b) assign the corresponding class labels. From the obtained persistent homology diagram in Figure 5a, we see that (a) there is one connected set and (b) that there is one hole, i.e., that there are two classes. We can further read off a separating distance of  $\approx 1.5$ .

Using this information, it is now possible to cluster the raw state and control trajectory data, cf. Figures 5b and 5c.

We investigate how separating the solutions into unimodal clusters impacts the warm-start strategies based on simplistic interpolation, and strategies relying on training generative models from the raw control trajectory samples. We now look at the overall mean and standard deviation of the control

trajectories and compare this to the mean of each of the two identified and labeled classes (cf. Figure 5d). As shown in the figure, if the multi-modality of the solutions is not taken into account, the information in the data effectively cancels out producing a mean trajectory (black) that encodes little useful structure and has a large standard deviation. This is due to the symmetry in the control trajectories—by summing them all terms vanish.

### C. Quadrotor navigation in a maze

In the previous experiment, we tested our methodology on a low-dimensional problem using a cartpole without considering environments with obstacles, where collision avoidance is needed. We now consider a set of experiments to investigate both the scalability in terms of problem dimensionality as well as to include multi-modality in the solutions introduced by collision avoidance. Additionally, we show the impact of the data separation on successfully learning initial seeds for the optimal control problem.

Quadrotors are agile multi-rotor drones that have become very popular in the past decade due to increased battery energy densities, more powerful motors, and reduced component prices. To model the quadrotor dynamics, we follow [41] with minor modifications: We do not consider the effects of air drag (or near ground effects) and control the rotor forces directly. Hence, the control inputs are  $\mathbf{u} \in \mathbb{R}^4$  (limited between 0 N to 5 N)



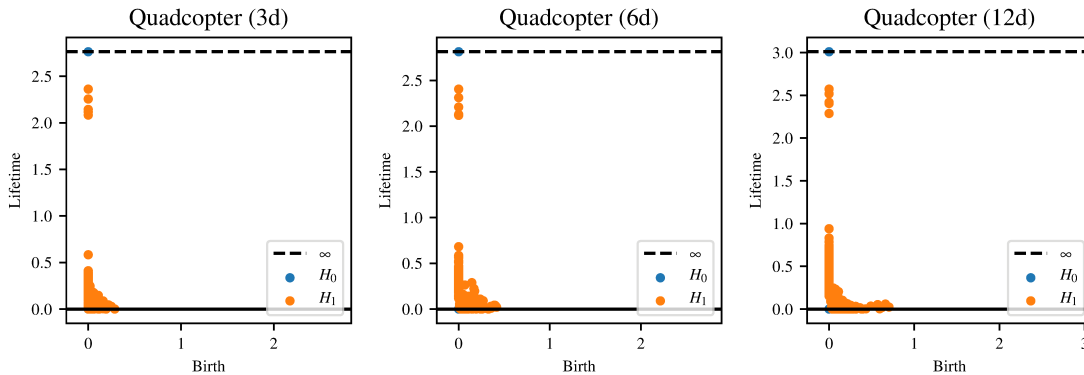


Figure 6. Filtration of the quadrotor dataset using different state spaces: translation only, translation and rotation, and full state space including velocities: The identified underlying topology and computation times are the same.

and the state is  $\mathbf{x} \in \mathbb{R}^{12}$  using Euler representation for the angular component of the floating base.

We sample a dataset of trajectories for a quadrotor flying from a start state uniformly sampled from  $[-3.25, -0.25]$  for each  $x$ ,  $y$ , and  $z$  to a goal position  $((x, y, z) = (1.75, 1.75, 1.75))$  while avoiding collision in an environment using the formulation described in [8], see Figure 1. The environment consists of three intersecting cylinders centered at the origin. We use  $T = 50$  knot points with  $\Delta t = 0.05$ s. We initialize the optimal control solver with a state trajectory obtained from RRT-Connect [42] and a control trajectory for hovering with noise sampled from the standard normal distribution ( $\mathcal{N}(0, 0.01)$ ) and solve until convergence.

Our dataset consists of 5535 valid state and control trajectories. We depict the persistent homology graphs in Figure 6 and automatically extract the number of clusters as six (five holes). Filtration on 3D (base position), 6D (base position and orientation), and 12D (base position, orientation, and spatial velocities) state spaces took 18.9 s, 18.5 s, and 22.0 s, respectively. Analogously to the cartpole example, we perform clustering using persistent homology. The labeled state trajectories are shown in Figure 1.

We now use the labeled samples to learn state and control predictors to generate initial seeds for our optimal control solver following Section IV. To benchmark, we randomly sample start states from the same range as in the dataset and predict state and control trajectories using the four comparison methods:

- 1) Cold-start: Controls initialized to achieve hovering and states initialized with the initial state for entire horizon.
- 2) MLP: Prediction from a single MLP each for both state and control trajectories (1-hidden layer, 200 neurons, 245,596 trainable parameters).
- 3) KNN-Regressor: Prediction from a KNN-Regressor method which averages over  $k$  neighbors.  $k \in [1, 10]$  has been brute-forced for best performance using a held-out validation set.
- 4) Our method: A Mixture-of-Experts setup trained on the separated data. Smaller MLP where trained for each of the continuous subsets and a gating network (with softmax activation) trained to select the most suitable expert (1-hidden layer, 50 neurons). The total trainable parameters across all experts and gating network are 252,472. Note,

this is the only class-aware prediction method in the test field.

We perform this benchmark for 500 random start states and show the mean and standard deviation of the convergence (cost evolution versus wall-clock time) in Figure 8. It is evident that the modality-aware method has a lower initial cost and faster convergence than other initialization methods. Note that the cold-start (no initial guess) performs better than a non-mode-aware learning method (MLP, KNN-Regressor). A likely explanation for this is that a roll-out of the predicted control trajectories has large dynamic defects compared with the predicted state trajectories. This can be seen by comparing the learned state initialization, the roll-out of the learned control initialization, and the final optimized trajectory in the accompanying video.<sup>5</sup> We have also evaluated the number of major iterations, final cost, and success rate as shown in Figure 7. The number of major iterations show the same trend as the convergence time in Figure 8, which is expected. The success rates for the four methods (cold-start, MLP, KNN regressor, and proposed) were 2.4%, 17.2%, 26.4%, and 99.8%, respectively.

#### D. Humanoid Reaching: Filtration in alternate state spaces

In addition to clustering initial seeds for dynamical optimal control, persistent homology can also be used as a tool to choose an alternate task space in which to represent the trajectory. We have illustrated this problem using the toy example in Section III-A where we talked about scaling of the velocity component, which is a trivial way of defining an alternate space. In more general scenarios, we look for spaces where the motion would produce simpler and more persistent topology. This is particularly relevant when we intend to track a kinematic trajectory using a Proportional-Derivative (PD) controller. A PD controller minimizes error in joint space. It is however very common to implement an operational space controller which is a PD controller minimizing error in task space. To ensure the controller can keep tracking the task, we want to choose a space in which the PD controller can track a kinematic reference trajectory by minimizing a Euclidean

<sup>5</sup><https://youtu.be/IUULTWCFxY8>

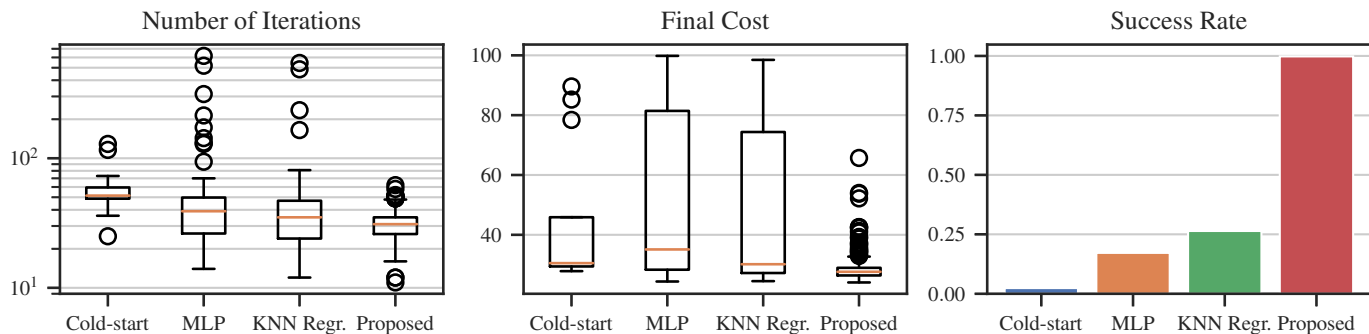


Figure 7. Warm-start comparison results on the quadrotor: Number of major iterations, final cost, and convergence success rate when initializing the optimal control problem using different initialization methods. We compute the number of iterations and final cost based on successful trials only.

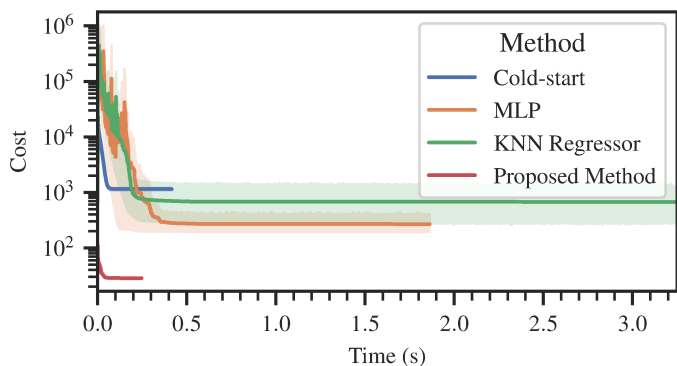


Figure 8. Cost evolution (mean, standard deviation) for solving the quadrotor task using BoxFDDP and state and control trajectory initializations predicted by different learning methods.

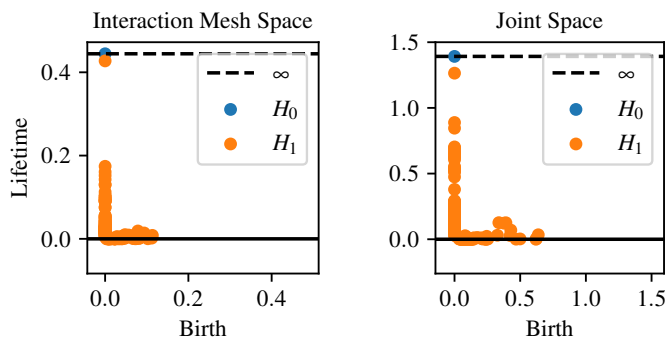


Figure 10. Persistence intervals for the interaction mesh space (left) and the joint space (right).

error. A space with a simpler topology would therefore be more suitable for tracking a reference trajectory.

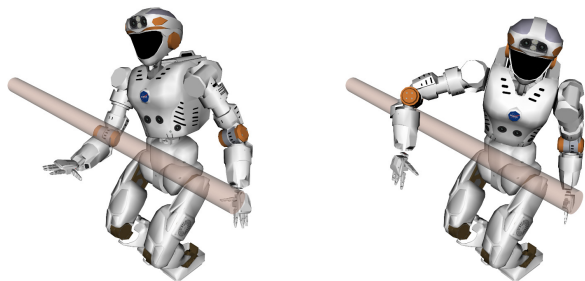


Figure 9. Valkyrie robot reaching scenario, showing two distinct reaching poses creating the bifurcation in the solution space.

We have generated a dataset of kinematic reference trajectories for the torso and arm of the Valkyrie humanoid robot (see Figure 9). We have planned 100 trajectories sampled at 200 knot points per trajectory. We have defined a reaching task with the reaching target in front of the robot but behind a horizontal bar. The task has two solutions: reaching above and below the bar. We have solved the problem using RRT-Connect to ensure that the trajectories are collision-free.

Figure 10 shows the results of filtration in joint space and an alternate, topological state space representation: the interaction

mesh [43]. The interaction mesh is hereby defined as the edges between points on the robot and key points in the environment capturing the relationship between the movement of the robot within its surroundings by computing the Laplace coordinates of these points. The coordinates in the interaction mesh space also have a much more uniform relative scale along each dimension which means that the persistent topology features are a lot more likely to emerge at the same scale across the whole space. The results show that the interaction mesh space representation has a much more clearly defined topology than the joint space. In the persistence diagram, this is visible as one  $H_1$  group far away from a cluster of short-lived noise. As a result, the topology of the trajectory data is more clearly represented in interaction mesh space, highlighting an interesting avenue for further work exploring the topology of other task spaces.

## VI. CONCLUSIONS

In this paper, we introduced a method to automatically identify the number of classes and cluster a dataset of trajectories using persistent homology. We then trained a generative model to compress and generalize the dataset for use as an initial seed in future optimizations. Our experiments explored this concept on optimal control problems involving dynamics as well as high-dimensional kinematic tasks and focused on establishing whether relevant topological information can be extracted to assist the encoding and learning stages of storing a memory-of-motion.

Our results confirm that considering the underlying topological features of the dataset is important and that tools from algebraic homology can be used to guide clustering. Indeed, exploring multiple modes can be important for warm-starting (cf. Figure 8) and further allows the development of ensemble methods that explore multiple warm-start guesses in parallel [10]. While these simultaneously explore multiple warm-starts, an alternative approach would be to test ranked warm-start modes/hypotheses akin to sequence prediction [6].

#### A. Limitations and Future work

One limitation of our approach is that the tools for filtration are very sensitive to the amount of data. Empirically, for a current laptop computer, distance matrices up to a few thousand rows/columns take tens of seconds to compute the persistent homology. However, it is important to note that, in practice, the time dimension can often be sub-sampled while preserving the salient features in the data. A key consideration here is that the cohomology computation can only capture features that are present in the data: If a topological hole is not captured in the discretized (or subsampled) data, a clustering method will not be able to extract it. A common example is thin obstacles for which trajectories may be collision-free at the discretized knot points but would require traversing through the obstacle to connect subsequent trajectory points. We have previously studied and proposed a solution to this challenge in [44]. Since our method proposes the use of a segment-to-segment distance, it permits the evaluation of distances in either discrete- or continuous-time allowing for these considerations to be incorporated. Additionally, the underlying structure of the dataset (i.e., number of classes) can be explored using subsets of the data without affecting the scalability of the clustering step—a feature we have explored and applied to scale to larger datasets. In this work, we did not explore homology groups in dimensions greater than 1. The  $H_2$  groups, in particular, can be used to cluster trajectories enclosing volumes of space, such as a sphere. We did not explore such motion here as it would require a large number of trajectories to produce a dense enough coverage for the  $H_2$  persistence to be reliably detected. This is an interesting future research direction, especially for systems and tasks with redundancy and periodicity.

The algorithms we used in this work for computing the homology groups did not take advantage of parallelism and available implementations are memory-intensive. Recent advances in computational homology have focused on leveraging massively parallel architectures to reduce the computation time by one to two orders of magnitude [45], [46]. We plan to evaluate and leverage these in future work.

One way to address dataset size could be to apply curve fitting techniques to reduce the dimension of the distance matrix prior to filtration. Curve fitting techniques and embeddings (e.g., splines or Bezier curves) can be used to represent longer time-series segments in place of the linear segments we have used in Section III-A. Alternatively, [47] previously applied sliding windows to discover periodicity in time-series data using persistent homology. It is worthwhile to explore whether a similar approach can be applied in our case instead of the dense filtration of the full dataset.

If the trajectories cover the solution space sparsely, the topology of the data may depend more on the random seed of the trajectory generator rather than the underlying solution space. This means that in some cases, a relatively small obstacle may be mistaken for a non-persistent  $H_1$  group and result in clustering the data incorrectly. This could lead to a warm-start trajectory passing through an obstacle. We could remedy this by either increasing the number of trajectories in the dataset, by utilizing a collision avoidance technique, or by artificially inflating the distance between a pair of trajectories passing around the obstacle. However, the former comes with increased computational cost while the latter is only possible if we know that the obstacle is present. Enforcing topology around known obstacles would be an interesting extension of this work.

State spaces including derived quantities such as velocities or accelerations pose a further challenge. To normalize/trade-off spaces with different units (i.e., scaling velocity down w.r.t. configuration), we apply intuition and manually scale each dimension. Latest work on multi-parameter persistent homology [48] offers new tools to compute the persistence of the trajectories with different amounts of scaling in a principled way. This problem could then be further expanded to analyzing the topology of arbitrary parameterized task spaces to discover spaces with a simpler topology that is favorable for simple operational space controllers.

Another challenge, also identified by others, is that we did not consider changing environments. Some authors [7], [9] outlined the possibility to enhance the problem parameterization with information on the size and location of geometric primitive shaped objects. We consider this inflexible, as fixed-size, basic representations always limit expressiveness and require vast numbers of samples to explore the increased dimensionality of the problem space. Instead, [5] investigated and compared learned situation descriptors. More recent work focused on learning latent space representations directly from sensor data such as point clouds [49]. Along with large, labeled datasets on 3D objects and shapes, these are promising avenues for further investigation.

Finally, we explored systems with continuous dynamics and straight-forward start and goal situations. It would be interesting to explore the use of tools from algebraic topology on tasks with discontinuous dynamics and periodicity, for instance, locomotion on legged platforms.

#### REFERENCES

- [1] C. Mastalli, R. Budhiraja, W. Merkt, G. Saurel, B. Hammoud, M. Naveau, J. Carpentier, L. Righetti, S. Vijayakumar, and N. Mansard, “Crocodyl: An Efficient and Versatile Framework for Multi-Contact Optimal Control,” in *Proc. IEEE Int. Conf. Rob. Autom. (ICRA)*, May 2020, pp. 2536–2542.
- [2] C. G. Atkeson and J. Morimoto, “Nonparametric Representation of Policies and Value Functions: A Trajectory-Based Approach,” in *Advances in Neural Information Processing Systems 15*, S. Becker, S. Thrun, and K. Obermayer, Eds. MIT Press, 2003, pp. 1643–1650.
- [3] M. Stolle and C. G. Atkeson, “Policies based on trajectory libraries,” in *Proc. IEEE Int. Conf. Rob. Autom. (ICRA)*, May 2006, pp. 3344–3349.
- [4] Y. Tassa, T. Erez, and W. D. Smart, “Receding Horizon Differential Dynamic Programming,” in *Advances in Neural Information Processing Systems 20*, J. C. Platt, D. Koller, Y. Singer, and S. T. Roweis, Eds. Curran Associates, Inc., 2008, pp. 1465–1472.
- [5] N. Jetchev and M. Toussaint, “Fast motion planning from experience: trajectory prediction for speeding up movement generation,” *Autonomous Robots*, vol. 34, no. 1, pp. 111–127, Jan 2013.

- [6] D. Dey, T. Y. Liu, M. Hebert, and J. A. Bagnell, "Contextual sequence prediction with application to control library optimization," in *Robotics: Science and Systems VIII*, N. Roy, P. Newman, and S. Srinivasa, Eds. MIT Press, 2013.
- [7] N. Mansard, A. Del Prete, M. Geisert, S. Tonneau, and O. Stasse, "Using a Memory of Motion to Efficiently Warm-Start a Nonlinear Predictive Controller," in *Proc. IEEE Int. Conf. Rob. Autom. (ICRA)*, May 2018, pp. 2986–2993.
- [8] W. Merkt, V. Ivan, and S. Vijayakumar, "Leveraging Precomputation with Problem Encoding for Warm-Starting Trajectory Optimization in Complex Environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Rob. Sys. (IROS)*, Oct 2018, pp. 5877–5884.
- [9] G. Tang and K. Hauser, "Discontinuity-Sensitive Optimal Control Learning by Mixture of Experts," in *Proc. IEEE Int. Conf. Rob. Autom. (ICRA)*, May 2019, pp. 7892–7898.
- [10] T. S. Lembono, A. Paolillo, E. Pignat, and S. Calinon, "Memory of Motion for Warm-Starting Trajectory Optimization," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2594–2601, April 2020.
- [11] S. Vijayakumar and S. Schaal, "Locally weighted projection regression: An O(n) algorithm for incremental real time learning in high dimensional space," in *Proceedings of the Seventeenth International Conference on Machine Learning (ICML)*, ser. ICML '00. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000, pp. 1079–1086.
- [12] C. K. I. Williams and C. E. Rasmussen, "Gaussian Processes for Regression," in *Advances in Neural Information Processing Systems 8*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds. MIT Press, 1996, pp. 514–520.
- [13] M. Toussaint and S. Vijayakumar, "Learning Discontinuities with Products-of-Sigmoids for Switching between Local Models," in *Proceedings of the 22nd International Conference on Machine Learning*, ser. ICML '05. New York, NY, USA: Association for Computing Machinery, 2005, pp. 904–911.
- [14] J. Aleotti and S. Caselli, "Robust trajectory learning and approximation for robot programming by demonstration," *Robotics and Autonomous Systems*, vol. 54, no. 5, pp. 409–413, 2006, the Social Mechanisms of Robot Programming from Demonstration.
- [15] T. Oates, L. Firoiu, and P. R. Cohen, "Clustering time series with hidden markov models and dynamic time warping," in *Proceedings of the IJCAI-99 workshop on neural, symbolic and reinforcement learning methods for sequence learning*. Citeseer, 1999, pp. 17–21.
- [16] K. Ogawara, J. Takamatsu, H. Kimura, and K. Ikeuchi, "Modeling manipulation interactions by hidden Markov models," in *Proc. IEEE/RSJ Int. Conf. Intell. Rob. Sys. (IROS)*, vol. 2, Sep. 2002, pp. 1096–1101 vol.2.
- [17] M. Stolle, H. Tappeiner, J. Chestnutt, and C. G. Atkeson, "Transfer of policies based on trajectory libraries," in *Proc. IEEE/RSJ Int. Conf. Intell. Rob. Sys. (IROS)*, Oct 2007, pp. 2981–2986.
- [18] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive Mixtures of Local Experts," *Neural Computation*, vol. 3, no. 1, pp. 79–87, 1991, pMID: 31141872.
- [19] G. E. Hinton, "Products of experts," *Proceedings of the 9th International Conference on Artificial Neural Networks*, pp. 1–6(5), 1999.
- [20] E. Pignat and S. Calinon, "Bayesian Gaussian Mixture Model for Robotic Policy Imitation," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4452–4458, Oct 2019.
- [21] C. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [22] F. T. Pokorný, M. Hawasly, and S. Ramamoorthy, "Topological trajectory classification with filtrations of simplicial complexes and persistent homology," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 204–223, 2016.
- [23] H. Edelsbrunner and J. Harer, "Persistent homology - a survey," *Discrete & Computational Geometry - DCG*, vol. 453, 01 2008.
- [24] G. Carlsson, "Topology and data," *Bulletin of the American Mathematical Society*, vol. 46, no. 2, pp. 255–308, 2009.
- [25] C. Chen and M. Kerber, "Persistent homology computation with a twist," in *Proceedings 27th European Workshop on Computational Geometry*, 2011.
- [26] N. Cavanna, M. Jahansair, and D. Sheehy, "A geometric perspective on sparse filtrations," in *Proceedings of the Canadian Conference on Computational Geometry (CCCG 2015)*, 06 2015.
- [27] C. Mastalli, W. Merkt, J. Marti-Saumell, H. Ferrolho, J. Sola, N. Mansard, and S. Vijayakumar, "A Direct-Indirect Hybridization Approach to Control-Limited DDP," 2020.
- [28] D. Q. Mayne, "A Second-order Gradient Method for Determining Optimal Trajectories of Non-linear Discrete-time Systems," *International Journal of Control*, vol. 3, no. 1, pp. 85–95, 1966.
- [29] D. H. Jacobson, "Differential dynamic programming methods for solving bang-bang control problems," *IEEE Transactions on Automatic Control*, vol. 13, no. 6, pp. 661–675, December 1968.
- [30] Y. Tassa, "Theory and Implementation of Biomimetic Motor Controllers," Ph.D. dissertation, Hebrew University of Jerusalem, 2011.
- [31] Y. Tassa, N. Mansard, and E. Todorov, "Control-limited differential dynamic programming," in *Proc. IEEE Int. Conf. Rob. Autom. (ICRA)*, May 2014, pp. 1168–1175.
- [32] E. W. Weisstein, "Homotopy," From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/Homotopy.html>, 1999–2019.
- [33] T. Kaczynski, K. Mischaikow, and M. Mrozek, *Computational homology*. Springer Science & Business Media, 2006, vol. 157.
- [34] R. Ghrist, "Barcodes: the persistent topology of data," *Bulletin of the American Mathematical Society*, vol. 45, no. 1, pp. 61–75, 2008.
- [35] V. J. Lumelsky, "On fast computation of distance between line segments," *Information Processing Letters*, vol. 21, no. 2, pp. 55–61, 1985.
- [36] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.
- [37] V. Ivan, Y. Yang, W. Merkt, M. P. Camilleri, and S. Vijayakumar, "EXOTica: An Extensible Optimization Toolset for Prototyping and Benchmarking Motion Planning and Control," in *Robot Operating System (ROS): The Complete Reference (Volume 3)*, A. Koubaa, Ed. Cham: Springer International Publishing, 2019, pp. 211–240.
- [38] U. Bauer, "Ripser: a lean C++ code for the computation of Vietoris–Rips persistence barcodes," <https://github.com/Ripser/ripser>, 2017.
- [39] C. Tralie, N. Saul, and R. Bar-On, "Ripser.py: A Lean Persistent Homology Library for Python," *The Journal of Open Source Software*, vol. 3, no. 29, p. 925, Sep 2018.
- [40] R. Tedrake, "Underactuated robotics: Algorithms for walking, running, swimming, flying, and manipulation (course notes for MIT 6.832)," <http://underactuated.mit.edu>, 2014–2019.
- [41] D. Mellinger, N. Michael, and V. Kumar, "Trajectory generation and control for precise aggressive maneuvers with quadrotors," *The International Journal of Robotics Research*, vol. 31, no. 5, pp. 664–674, 2012.
- [42] J. J. Kuffner and S. M. LaValle, "RRT-Connect: An efficient approach to single-query path planning," in *Proc. IEEE Int. Conf. Rob. Autom. (ICRA)*, vol. 2. IEEE, 2000, pp. 995–1001.
- [43] V. Ivan, D. Zarubin, M. Toussaint, T. Komura, and S. Vijayakumar, "Topology-based representations for motion planning and generalization in dynamic environments with interactions," *The International Journal of Robotics Research*, vol. 32, no. 9-10, pp. 1151–1163, 2013.
- [44] W. Merkt, V. Ivan, and S. Vijayakumar, "Continuous-Time Collision Avoidance for Trajectory Optimization in Dynamic Environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Rob. Sys. (IROS)*, Nov 2019, pp. 7248–7255.
- [45] S. Zhang, M. Xiao, and H. Wang, "GPU-Accelerated Computation of Vietoris-Rips Persistence Barcodes," in *36th International Symposium on Computational Geometry (SoCG 2020)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), S. Cabello and D. Z. Chen, Eds., vol. 164. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020, pp. 70:1–70:17. [Online]. Available: <https://drops.dagstuhl.de/opus/volltexte/2020/12228>
- [46] D. Morozov and A. Nigmatov, "Towards lockfree persistent homology," in *Proceedings of the 32nd ACM Symposium on Parallelism in Algorithms and Architectures*, 2020, pp. 555–557.
- [47] J. A. Perea and J. Harer, "Sliding Windows and Persistence: An Application of Topological Methods to Signal Analysis," *Foundations of Computational Mathematics*, vol. 15, no. 3, pp. 799–838, Jun 2015.
- [48] H. A. Harrington, N. Otter, H. Schenck, and U. Tillmann, "Stratifying Multiparameter Persistent Homology," *SIAM Journal on Applied Algebra and Geometry*, vol. 3, no. 3, pp. 439–471, 2019.
- [49] A. H. Qureshi, Y. Miao, A. Simeonov, and M. C. Yip, "Motion Planning Networks: Bridging the Gap Between Learning-Based and Classical Motion Planners," *IEEE Transactions on Robotics*, vol. 37, no. 1, pp. 48–66, Feb 2021.



**Wolfgang Merkt** received the B.Eng.(Hons) degree in mechanical engineering with management and the M.Sc.(R) and Ph.D. degrees in robotics and autonomous systems from the University of Edinburgh in 2014, 2015, and 2019, respectively. He is currently a postdoctoral researcher at the Oxford Robotics Institute, University of Oxford. During his Ph.D., he worked on trajectory optimization and warm-starting optimal control for high-dimensional systems and humanoid robots. Wolfgang's research interests include fast optimization-based methods for planning and control, loco-manipulation, and legged robots.



**Sethu Vijayakumar** is the Professor of Robotics at The University of Edinburgh, U.K., an Adjunct Faculty with the University of Southern California, Los Angeles, CA, USA and the Founding Director of the Edinburgh Centre for Robotics. His research interests include statistical machine learning, anthropomorphic robotics, planning, multi-objective optimization and optimal control in autonomous systems as well as the study of human motor control. Prof. Vijayakumar helps shape and drive the national Robotics and Autonomous Systems (RAS) agenda in his role as the Programme co-Director for Artificial Intelligence (AI) with The Alan Turing Institute, the UK's national institute for data science and AI. He is a Fellow of the Royal Society of Edinburgh, a Judge on BBC Robot Wars and winner of the 2015 Tam Dalyell Prize for excellence in engaging the public with science.



**Vladimir Ivan** received his Ph.D. on the topic of Motion synthesis in topology-based representations at the University of Edinburgh where he is currently working as a Senior Researcher in the School of Informatics. He has previously received a M.Sc. in Artificial Intelligence specializing in Intelligent Robotics at the University of Edinburgh and a B.Sc. in AI and Robotics from the University of Bedfordshire. Vladimir's research interests are motion planning and modeling, topology, humanoid and legged robotics, space robotics, and shared autonomy.



**Traiko Dinev** received a B.Sc. degree in computer science (with electronics) and an M.Sc. in robotics and autonomous systems from the University of Edinburgh. He is currently pursuing a PhD degree with the Edinburgh Centre for Robotics at the University of Edinburgh. Traiko's research interests include motion planning for dynamic locomotion, hardware-software co-design for planning and control, and applied statistical machine learning.



**Ioannis Havoutis** is a Lecturer in Robotics at the University of Oxford. He is part of the Oxford Robotics Institute and a co-lead of the Dynamic Robot Systems group. His focus is on approaches for dynamic whole-body motion planning and control for legged robots in challenging domains. From 2015 to 2017, he was a postdoc at the Robot Learning and Interaction Group at the Idiap Research Institute. Previously, from 2011 to 2015, he was a senior postdoc at the Dynamic Legged System lab at the Istituto Italiano di Tecnologia. He holds a Ph.D. and M.Sc. from the University of Edinburgh.